

DESI-Konsortium [Hrsg.]

Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie

Weinheim u.a. : Beltz 2008, 459 S. - (Beltz Pädagogik)



Quellenangabe/ Reference:

DESI-Konsortium [Hrsg.]: Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim u.a. : Beltz 2008, 459 S. - (Beltz Pädagogik) - URN: urn:nbn:de:0111-opus-31494 - DOI: 10.25656/01:3149

<https://nbn-resolving.org/urn:nbn:de:0111-opus-31494>

<https://doi.org/10.25656/01:3149>

in Kooperation mit / in cooperation with:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

DESI-Konsortium (Hrsg.)

Unterricht und Kompetenzerwerb in Deutsch und Englisch

Ergebnisse der DESI-Studie

Beltz Verlag · Weinheim und Basel

Diese Studie wurde im Auftrag der Kultusministerkonferenz erstellt.
Für die Richtigkeit der Ergebnisse der Studie tragen die Herausgeber
die Verantwortung.

Herausgeber:

Eckhard Klieme (Sprecher des DESI-Konsortiums), Wolfgang Eichler,
Andreas Helmke, Rainer H. Lehmann, Günter Nold, Hans-Günter Rolff,
Konrad Schröder, Günther Thomé und Heiner Willenberg.

Das Werk und seine Teile sind urheberrechtlich geschützt.
Jede Nutzung in anderen als den gesetzlich zugelassenen Fällen
bedarf der vorherigen schriftlichen Einwilligung des Verlages.
Hinweis zu § 52a UrhG: Weder das Werk noch seine Teile dürfen
ohne eine solche Einwilligung eingescannt und in ein Netzwerk
eingestellt werden. Dies gilt auch für Intranets von Schulen
und sonstigen Bildungseinrichtungen.

© 2008 Beltz Verlag · Weinheim und Basel
www.beltz.de

Herstellung: Klaus Kaltenberg

Satz: Deutsches Institut für Internationale Pädagogische Forschung

Druck: Druck Partner Rübelmann, Hemsbach

Printed in Germany

ISBN 978-3-407-25491-7

Inhaltsverzeichnis

Konzeption der Studie

Eckhard Klieme

1	Systemmonitoring für den Sprachunterricht.....	1
---	--	---

Bärbel Beck / Svenja Bundt / Jens Gomolka

2	Ziele und Anlage der DESI-Studie	11
---	--	----

Thamar Dubberke / Birgit Harks

2.5	Zur curricularen Validität der DESI-Aufgaben: Ergebnisse eines Expertenratings.....	26
-----	--	----

Johannes Hartig / Nina Jude / Wolfgang Wagner

3	Methodische Grundlagen der Messung und Erklärung sprachlicher Kompetenzen	34
---	--	----

Hans-Günter Rolff / Jan von der Gathen

4	Rückmeldungen an Lehrkräfte und Rezeption.....	55
---	--	----

Leistungsverteilungen im Deutschen und Englischen

Steffen Gailberger / Heiner Willenberg

5	Leseverstehen Deutsch	60
---	-----------------------------	----

Heiner Willenberg

6	Wortschatz Deutsch	72
---	--------------------------	----

Michael Krelle / Heiner Willenberg

7	Argumentation Deutsch	81
---	-----------------------------	----

Astrid Neumann / Rainer H. Lehmann

8	Schreiben Deutsch.....	89
---	------------------------	----

Günther Thomé / Wolfgang Eichler

9	Rechtschreiben Deutsch.....	104
---	-----------------------------	-----

Wolfgang Eichler

10	Sprachbewusstheit Deutsch	112
----	---------------------------------	-----

Günter Nold / Henning Rossa

11	Hörverstehen Englisch.....	120
----	----------------------------	-----

Günter Nold / Henning Rossa / Kyriaki Chatzivassiliadou

12	Leseverstehen Englisch.....	130
----	-----------------------------	-----

	<i>Claudia Harsch / Konrad Schröder / Astrid Neumann</i>	
13	Schreiben Englisch	139
	<i>Claudia Harsch / Konrad Schröder</i>	
14	Textrekonstruktion Englisch	149
	<i>Günter Nold / Henning Rossa</i>	
15	Sprachbewusstheit Englisch	157
	<i>Günter Nold / Henning Rossa</i>	
16	Sprechen Englisch	170
	<i>Hermann-Günter Hesse / Kerstin Göbel / Nina Jude</i>	
17	Interkulturelle Kompetenz	180

Individuelle und familiale Bedingungsfaktoren sprachlicher Kompetenzen

	<i>Nina Jude / Eckhard Klieme / Wolfgang Eichler / Rainer H. Lehmann / Günter Nold / Konrad Schröder / Günther Thomé / Heiner Willenberg</i>	
18	Strukturen sprachlicher Kompetenzen	191
	<i>Johannes Hartig / Nina Jude</i>	
19	Sprachkompetenzen von Mädchen und Jungen	202
	<i>Hermann-Günter Hesse / Kerstin Göbel / Johannes Hartig</i>	
20	Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache	208
	<i>Wolfgang Wagner / Andreas Helmke / Friedrich-Wilhelm Schrader / Wolfgang Eichler / Günther Thomé / Heiner Willenberg</i>	
21	Selbstkonzept und Motivation im Fach Deutsch	231
	<i>Andreas Helmke / Friedrich-Wilhelm Schrader / Wolfgang Wagner / Günter Nold / Konrad Schröder</i>	
22	Selbstkonzept, Motivation und Englischleistung	244
	<i>Friedrich-Wilhelm Schrader / Andreas Helmke / Wolfgang Wagner / Wolfgang Eichler / Günther Thomé / Heiner Willenberg</i>	
23	Lernstrategien im Fach Deutsch	258
	<i>Friedrich-Wilhelm Schrader / Andreas Helmke / Wolfgang Wagner / Günter Nold / Konrad Schröder</i>	
24	Lernstrategien im Fach Englisch	270
	<i>Hans-Günter Rolf / Michael Leucht / Ernst Rösner</i>	
25	Sozialer und familialer Hintergrund	283

Unterricht und Lehrerkompetenzen

Andreas Helmke / Eckhard Klieme

- 26 Unterricht und Entwicklung sprachlicher Kompetenzen..... 301

*Holger Ehlers / Nina Jude / Eckhard Klieme / Andreas Helmke /
Wolfgang Eichler / Heiner Willenberg*

- 27 Soziodemografische und fachdidaktisch relevante Merkmale
von Deutsch-Lehrpersonen..... 313

*Eckhard Klieme / Nina Jude / Dominique Rauch / Holger Ehlers /
Andreas Helmke / Wolfgang Eichler / Günther Thomé / Heiner Willenberg*

- 28 Alltagspraxis, Qualität und Wirksamkeit des
Deutschunterrichts 319

*Tuyet Helmke / Andreas Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Günter Nold / Konrad Schröder*

- 29 Die Videostudie des Englischunterrichts..... 345

*Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Günter Nold / Konrad Schröder*

- 30 Soziodemografische und fachdidaktisch relevante Merkmale
von Englischlehrpersonen..... 364

*Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Günter Nold / Konrad Schröder*

- 31 Alltagspraxis des Englischunterrichts 371

*Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Eckhard Klieme / Günter Nold / Konrad Schröder*

- 32 Wirksamkeit des Englischunterrichts 382

Kerstin Göbel / Hermann-Günter Hesse

- 33 Vermittlung interkultureller Kompetenzen im
Englischunterricht 398

Institutionelle Bedingungsfaktoren sprachlicher Kompetenzen

Brigitte Steinert / Johannes Hartig / Eckhard Klieme

- 34 Institutionelle Bedingungen der Sprachkompetenzen..... 411

Günter Nold / Johannes Hartig / Silke Hinz / Henning Rossa

- 35 Klassen mit bilinguaalem Sachfachunterricht: Englisch als
Arbeitssprache 451

- Die Autorinnen und Autoren..... 458

Eckhard Klieme

1 Systemmonitoring für den Sprachunterricht

Die Studie „Deutsch Englisch Schülerleistungen International“ (DESI) untersuchte die sprachlichen Leistungen der Schülerinnen, den Unterricht sowie professionelle Leitvorstellungen und kooperative Praktiken der Fachkollegien in den Fächern Deutsch und Englisch. Die Untersuchung wurde im Jahr 2001 von der Kultusministerkonferenz (KMK) als erste nationale Schulleistungsstudie in Auftrag gegeben. Unter Federführung des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) wurden dafür von einem Konsortium aus Bildungsforschern und Fachdidaktikern neue Testverfahren entwickelt. Etwa 11.000 Schülerinnen und Schüler der neunten Jahrgangsstufe aller Schularten wurden zu Beginn und am Ende des Schuljahres 2003/04 befragt und getestet; hinzu kamen Befragungen von Lehrkräften, Eltern und Schulleitungen sowie Videoaufnahmen im Englischunterricht. Als bundesweit repräsentative Untersuchung und durch ihre breitgefächerte Anlage ermöglicht die Studie differenzierte Aussagen über Lehr-Lern-Prozesse und den Erwerb sprachlicher Kompetenzen, die für Unterrichtspraxis, Lehrerbildung und Bildungspolitik gleichermaßen wichtig sind. Die zentralen Befunde wurden im März 2006, unmittelbar nach Abschluss des Projekts, der KMK vorgestellt und der Öffentlichkeit auf der Homepage des DIPF (www.dipf.de/desi) zugänglich gemacht. Die theoretische und methodische Basis dieser Arbeiten hat das Konsortium in der Folgezeit weiter ausgebaut und um detaillierte Informationen ergänzt. Die Ergebnisse der Forschungsarbeit werden im vorliegenden Band veröffentlicht.

1.1 Zielsetzung und Anspruch von DESI

DESI sollte einerseits den Leistungsstand der Schülerinnen und Schüler der neunten Jahrgangsstufe im Englischen und im Deutschen erfassen und andererseits Anregungen für die Revision von Curricula, für Lehrtexte und Unterrichtsmaterialien, für die Lehreraus- und -fortbildung und vor allem für die Unterrichtsgestaltung geben.

DESI zielt somit – anders als die international eingebundenen Studien wie TIMSS und PISA, anders auch als der Ländervergleich PISA-E – nicht primär auf Leistungsvergleiche ab, seien sie international oder intranational. Ein „Ranking“ der Länder der Bundesrepublik Deutschland war im Rahmen der DESI-Studie zu keinem Zeitpunkt vorgesehen. Das „I“ im Projektnamen steht für die internationale Anschlussfähigkeit der Studie; gleichwohl war dies nicht Gegenstand des Arbeitsauftrags der KMK. Realisiert wurde bislang eine mit DESI-Instrumenten durchgeführte Untersuchung der Deutsch-Kompetenzen von Schülerinnen und Schülern an den deutschen Schulen in Südtirol (vgl. Beck/Dahl 2006).

Da DESI eine Stichprobenuntersuchung ist, steht auch die vergleichende Rückmeldung an die Schulen nicht im Zentrum – anders als etwa in den neueren Vergleichsarbeiten auf Länderebene. Die Aussagekraft von DESI gründet sich auf das wissenschaftliche Design der Tests und Auswertungen: (a) die Beschreibung von Kompetenzstrukturen im Deutschen und im Englischen, (b) die kriteriumsorientierte Diagnose des Leistungsstands der Schülerinnen und Schüler auf nationaler Ebene sowie (c) die Analyse von Bedingungen des Leistungszuwachses, mit Schwerpunkten bei Unterricht, Schulqualität und familiären Voraussetzungen.

Differenzierte Erfassung von Sprachkompetenzen

Der besondere Charakter von DESI wurde durch die dem Projekt vorausgegangenen Diskussionen im Rahmen der Kultusministerkonferenz und ihrer Gremien deutlich. Im Anschluss an den Konstanzer Beschluss der KMK vom Oktober 1997 hatten sich die Länder dafür ausgesprochen, eine nationale Ergänzung zu TIMSS und PISA auszuschreiben, um die Erkenntnisse der internationalen Projekte zu ergänzen und Lücken in der Bandbreite der erfassten Leistungsbereiche für den nationalen Rahmen auszugleichen. Insbesondere die Erfassung der Fremdsprachenkompetenz unter Einbeziehung der aktiven Sprachbeherrschung, des Hörverstehens und der kommunikativen Kompetenzen in Aktion und Reaktion sollte das angestrebte Projekt gewährleisten. Ein besonderes Interesse galt dem Englischen als einem bedeutsamen Mittel für die internationale Kommunikation; daher wurden über die unmittelbaren Sprachkenntnisse hinaus auch interkulturelle Kommunikationsfähigkeiten in das Untersuchungsprogramm aufgenommen. Ebenso differenziert sollten die Kompetenzen im aktiven und passiven Gebrauch der Verkehrssprache Deutsch untersucht werden. Ergänzend zur Lesekompetenz, die als eine eher fächerübergreifende Domäne auch in den PISA-Studien erfasst wird, sollten der reflexive Umgang mit sprachlichen Strukturen, der schriftsprachliche Ausdruck, Aspekte von Kommunikations- und Argumentationsfähigkeit erfasst sowie Basisfähigkeiten im Wortschatz und in der Rechtschreibung beurteilt werden. Gerade aus der Verbindung der beiden Sprachen können sich bedeutsame Hinweise zur Erklärung von Leistungsunterschieden und für zu entwickelnde Handlungsstrategien ergeben. Dies bezieht sich auf den Zusammenhang zwischen der Entwicklung der Erstsprache und dem Erwerb einer ersten Fremdsprache, aber auch auf sprachliche Kompetenzen von Schülerinnen und Schülern, für die Deutsch nicht die Erstsprache ist.

DESI hat für das Fach Deutsch und für das Fach Englisch insgesamt 15 unterschiedliche Kompetenzen beschrieben (vgl. Beck/Klieme 2007). In intensiver Zusammenarbeit zwischen Fachdidaktikern, Schulforschern und Testexperten ist es gelungen, diese Vielfalt von Zielbereichen in Aufgaben und schließlich in Kompetenzmodelle umzusetzen. In DESI wurde exemplarisch eine gleichermaßen fachdidaktisch und psychometrisch fundierte, interdisziplinäre Arbeit an Kompetenzmodellen und Leistungstests geleistet. Diese Erfahrungen sind bereits in die Arbeiten der Kultusministerkonferenz zur Entwicklung und Normierung

von Bildungsstandards (Kultusministerkonferenz, 2005, www.kmk.org) eingeflossen. Die Englisch-Tests wurden mit Blick auf den „Gemeinsamen Europäischen Referenzrahmen für Sprachen“ (GERS; Goethe Institut Inter Nationes u.a., 2001) entwickelt und geben als erstes Verfahren eine empirisch abgesicherte Rückmeldung darüber, inwieweit Schülerinnen und Schüler in Deutschland in der Lage sind, derartige Standards tatsächlich zu erreichen. Die wissenschaftlichen Arbeiten im Rahmen des DESI-Projekts sollen darüber hinaus in die laufenden Bemühungen um einen Vergleich sprachlicher Kompetenzen in Europa mit einfließen und haben bereits internationale Beachtung gefunden.

Die Testentwicklung für DESI stützte sich auf detaillierte Lehrplananalysen für alle Länder und Schularten. Die Studie ist daher in der Art der Anforderungen und im Anspruchsniveau besonders eng auf die curricularen Anforderungen abgestimmt. Angesichts des hohen Anspruchsniveaus der Lehrpläne sind die Kompetenzmodelle im oberen Leistungsbereich stärker ausdifferenziert als im unteren. Aufgrund dieser curricularen Orientierung und der differenzierten Erfassung von Teilkompetenzen sind die DESI-Kompetenzmodelle mit den Kompetenzmodellen von PISA nicht vergleichbar. Aus den wissenschaftlichen Erfahrungen der DESI-Studie lässt sich ableiten, dass eine zentrale Herausforderung für die Normierung der KMK-Bildungsstandards darin liegt, Kompetenzmodelle auch im unteren Leistungsbereich auszdifferenzieren.

Erklärung unterschiedlicher Leistungsentwicklung

Ein Ziel von DESI war es, Erklärungsansätzen für unterschiedliche Leistungsniveaus bzw. unterschiedliche Lernzuwächse im Verlauf der neunten Jahrgangsstufe nachzugehen. Damit war die Bearbeitung von Fragestellungen verlangt, die sich im Kern auf fachdidaktische Konzepte stützen, aber auch darüber hinausgehen, denn unterrichtliche, schulische und außerschulische Einflüsse müssen gemeinsam untersucht und in ihrer Interdependenz verstanden werden. Die Berücksichtigung sozialer Kontextbedingungen, Fragen der Schulqualität, Bedingungen des individuellen wie des institutionellen Lernens, Lerngelegenheiten innerhalb und außerhalb der Schule, Sprachbiographie und innerfamiliäre Bedingungen des Kompetenzerwerbs waren in die Projektkonzeption einzubinden. Faktoren wie kognitive Lernvoraussetzungen, Motivation und außerschulische Lernumwelten bestimmen mit, was gelernt und letztlich, welcher Ertrag im Bildungswesen erzielt wird. Steuernde Einflussnahmen auf den Unterricht allein, ohne die sozialen und individuellen Eigenheiten der Schüler ins Kalkül zu ziehen, würden zu kurz greifen.

Die anspruchsvollen Erwartungen an das Projekt erfordern eine komplexe Anlage der Untersuchung. Die Fragen nach der Optimierung des Unterrichts und der Bildungsmaßnahmen erfordern prozessorientierte Forschungen, die allein mit Tests und Fragebögen nicht zu leisten sind, weshalb im Englischen eine videogestützte Unterrichtsanalyse vorgenommen wurde. Die Frage nach der Effizienzsteigerung schulischer Maßnahmen impliziert die Erklärung von Leistungszuwächsen, die nur

über die Untersuchung von Kompetenzen im Längsschnitt zu leisten ist. Die Suche nach zweckmäßigen Maßnahmen zur Verbesserung des Bildungssystems erfordert komplexe, mehrbenenanalytische Auswertungsdesigns.

Die Analysen sind in ein Rahmenmodell eingebettet, das als Angebot-Nutzungs-Modells formuliert ist (vgl. Kap. 26). In Bezug auf die „Produktion“ von Schulleistungen wird der Unterricht als ein in sich hoch strukturiertes Angebot betrachtet, dessen Qualität und Quantität für das Lernergebnis sehr wichtig sind, es aber keineswegs determinieren. Schule und Unterricht können nur Lerngelegenheiten bereitstellen, die von den Lernenden genutzt werden müssen. Die Schüler nehmen das Angebot nicht einfach direkt und quasi mechanisch auf, sondern gestalten konstruktiv die Lernsituation mit und entwickeln dabei Wissen und Kompetenzen. Ob, wie effizient und wie nachhaltig dies erfolgt, hängt vom Lernpotenzial der Schüler, vom Kontext und von verschiedenen Unterrichtsmerkmalen ab. Die Frage, die DESI in diesem Zusammenhang zu beantworten sucht, ist welche Merkmale von Lehrpersonen und des Fachunterrichts für den Kompetenzzuwachs eine Rolle spielen. Auf Seiten der Lehrpersonen standen Merkmale, denen aus unterrichtspsychologischer und fachdidaktischer Sicht eine lern- und motivationsförderliche Rolle zugeschrieben wird, im Mittelpunkt der Untersuchung. Für den Unterricht selbst wird in DESI die Qualität von Unterrichtsprozessen von der Qualität des Lehr-Lern-Materials und der Quantität des Unterrichts, also der Lernzeit, unterschieden. Im Zentrum steht dabei die Identifikation von übergeordneten Faktoren der Qualität des Fachunterrichts wie z.B. schülerorientiertes Unterrichtsklima, kognitive Aktivierung und effiziente Klassenführung.

1.2 Aufbau dieses Bandes

Der hier vorliegende Band geht in einem einleitenden Teil auf Ziele und Anlage der DESI Studie (Kap. 2), die methodischen Grundlagen der Sprachkompetenzmessung und Erklärung (Kap. 3), sowie die Rückmeldung der DESI Ergebnisse an die Lehrkräfte (Kap. 4) ein. Die darauf folgenden Kapitel setzen die Messung von sprachlichen Kompetenzen, wie sie bei Beck und Klieme (2007) dargestellt wurden, voraus und behandeln die für die Unterrichtspraxis relevanten Fragen in vier übergeordneten Themenblöcken. Im ersten Teil werden die Leistungsverteilungen im Deutschen und Englischen dargestellt. Dies gewährt Einblick in die sprachlichen Kompetenzen von Schülerinnen und Schülern aller Bildungsgänge am Ende der neunten Jahrgangstufe und dient als Grundlage für spätere Wirkungsanalysen. Die drei darauffolgenden Teile des Buchs bringen Bedingungsfaktoren auf verschiedenen Ebenen in Zusammenhang mit sprachlichen Kompetenzen. Zunächst werden auf Ebene des Individuums die individuellen und familialen Bedingungsfaktoren sprachlicher Kompetenzen analysiert, anschließend wird auf Ebene der Klasse der Einfluss des Unterrichts und der Lehrkräfte untersucht und schließlich auf Ebene der Schule nach institutionellen Bedingungsfaktoren sprachlicher Kompetenzen geforscht.

Leistungsverteilungen im Deutschen und Englischen

In den Kapiteln 5 bis 10 werden die Leistungsverteilungen im Deutschen in den Kompetenzbereichen Leseverstehen (Kap. 5), Wortschatz (Kap. 6), Argumentation (Kap. 7), Schreiben (Kap. 8), Rechtschreibung (Kap. 9) und Sprachbewusstheit (Kap. 10) vorgestellt; in den Kapiteln 11 bis 17 folgen die Darstellungen der Leistungsverteilungen im Englischen in den Kompetenzbereichen Hörverstehen (Kap. 11), Leseverstehen (Kap. 12), Schreiben (Kap. 13), Textrekonstruktion (Kap. 14), Sprachbewusstheit (Kap. 15), Sprechen (Kap. 16) und interkulturelle Kompetenz (Kap. 17). Einleitend wird dabei zunächst in aller Kürze auf das jeweilige Kompetenzmodell eingegangen. Für ein tieferes Verständnis muss an dieser Stelle auf den ersten Berichtsband (Beck/Klieme 2007) verwiesen werden. Die Operationalisierung wird auch im vorliegenden Buch anhand von Beispielaufgaben erläutert, bevor die Kompetenzniveaus definiert werden. Daran schließt sich als Kernstück eine kommentierte Darstellung der Leistungsverteilung in Jahrgangsstufe neun an, sowohl über die Bildungsgänge hinweg als auch nach Bildungsgang differenziert. Für die Kompetenzbereiche, die am Anfang und am Ende der neunten Jahrgangsstufe erhoben wurden, wird zudem die Kompetenzentwicklung innerhalb des Schuljahrs vorgestellt. Einige wichtige differentielle Effekte, wie Geschlechtsunterschiede im jeweiligen Kompetenzbereich und der Effekt der Erstsprache auf die Leistung werden zusätzlich dargestellt und jedes Kapitel schließt mit Schlussfolgerungen für die Fachdidaktik und den Fachunterricht.

Individuelle und familiale Bedingungsfaktoren sprachlicher Kompetenzen

Kapitel 18 befasst sich mit Strukturen sprachlicher Kompetenzen und bringt die in den Kapiteln 5 bis 17 dargestellten sprachlichen Kompetenzbereiche theoriebasiert und empirisch in einen Zusammenhang. Es bildet den Übergang von dem auf einzelne Kompetenzen fokussierenden ersten Teil des Buchs zu den die Bedingungen sprachlicher Kompetenzen auf verschiedenen Ebenen erforschenden folgenden Abschnitten. Die Kapitel 19 bis 25 befassen sich mit individuellen und familialen Bedingungsfaktoren sprachlicher Kompetenzen. Merkmale der Schülerinnen und Schüler sowie Merkmale der Herkunftsfamilie, dem Ort der ersten – auch sprachlichen – Sozialisation, werden mit sprachlichen Kompetenzen in Verbindung gebracht. Dabei befassen sich die Kapitel zu Sprachkompetenzen von Jungen und Mädchen (Kap. 19), zu Sprachkompetenzen von Schülern nicht deutscher Herkunftssprache (Kap. 20) und dem sozialen und familialen Hintergrund (Kap. 25) mit Merkmalen, die dem Unterricht vorgelagert sind. Das Selbstkonzept und die Motivation in den Fächern Deutsch (Kap. 21) und Englisch (Kap. 22) hingegen sind Teil des Gestaltungsraums von Unterricht, wie auch die Lernstrategien der Schülerinnen und Schüler in Deutsch (Kap. 23) und Englisch (Kap. 24); diese Merkmale können somit sowohl als Bedingungen wie auch als Ergebnisse des Unterrichts verstanden werden.

Unterricht und Lehrerkompetenzen

Eine zentrale Fragestellung von DESI ist, welche Merkmale von Lehrpersonen und des Fachunterrichts für den Kompetenzzuwachs eine Rolle spielen. Nachdem Leistungsverteilungen und individuelle Bedingungen sprachlicher Kompetenzen behandelt wurden, sind deshalb der Unterricht und die Lehrer Gegenstand der folgenden Kapitel. In Kapitel 26 wird ein theoretisches Modell für den Einfluss des Unterrichts auf die Entwicklung sprachlicher Kompetenzen vorgestellt und es wird die Analysestrategie erläutert, nach der in den Kapiteln 28 und 32 die Wirksamkeit bestimmter Merkmale des Unterrichts erforscht wird. Die Informationen über den Unterricht stammen jeweils aus der Befragung sowohl der Schülerinnen und Schüler als auch der Fachlehrkräfte, wodurch der Zugriff auf Unterricht aus beiden Perspektiven erfolgt. Für den Englischunterricht stehen zudem durch die Videostudie (Kap. 29) Beobachtungsdaten zum Englischunterricht zur Verfügung. Über diese zentralen Wirksamkeitsanalysen hinaus werden die Alltagspraxis des Deutsch- (Kap. 28) und Englischunterrichts (Kap. 31), sowie die für den Unterricht relevanten Merkmale der Deutsch- (Kap. 27) und Englischlehrer (Kap. 30) dargestellt. Wie schon bei der Darstellung der Leistung in unterschiedlichen Kompetenzbereichen wird auch hier gesondert auf die Vermittlung interkultureller Kompetenzen (Kap. 33) eingegangen.

Institutionelle Bedingungsfaktoren sprachlicher Kompetenzen

Zusätzlich zum Unterricht bezieht DESI darüber hinausgehende institutionelle Bedingungsfaktoren sprachlicher Kompetenzen ein (Kap. 34). Hierbei stützen sich die Analysen auf Befragungen der Fachkollegien und der Schulleitungen zu Fragen der pädagogischen Zielsetzung und Kooperationsstrukturen. Dem bilingualen Sachfachunterricht, der als eine der zentralen Innovationen für den Fremdsprachenunterricht gilt, ist zudem ein eigenständiges Kapitel (Kap. 35) am Ende des Buchs gewidmet.

1.3 DESI als Vorreiter für die Arbeit an Bildungsstandards und Kompetenzmodellen

Mit seinem Schwerpunkt auf fachdidaktisch und empirisch-pädagogisch fundierter Testentwicklung hat DESI bereits wichtige Beiträge zur aktuellen Diskussion in Bildungsforschung und Bildungspolitik geleistet. Diese Diskussion ist gegenwärtig u.a. durch die Arbeit an den Bildungsstandards (Kultusministerkonferenz, 2005) geprägt. Arbeitsgruppen der KMK haben curriculare Standards für die Leistungen der Schülerinnen und Schüler am Ende der Primarstufe und der Sekundarstufe I entwickelt, unter anderem im Deutschen und in der ersten Fremdsprache. Erklärtes Ziel der KMK war dabei die Ausarbeitung von Kompetenzmodellen, die systematisch die Teilkompetenzen innerhalb eines Faches und deren Abstufungen beschrei-

ben, und – damit verbunden – die Entwicklung und Skalierung von Tests, mit denen die Erreichung von Standards überprüft werden kann. Verantwortlich ist hierfür inzwischen das Institut für Qualitätsentwicklung im Bildungswesen (IQB, Rupp u.a. im Druck), das die Bildungsstandards für Deutsch und die erste Fremdsprache 2008 deutschlandweit in der Sekundarstufe I normieren wird. In 2009 wird als Ersatz für die früheren PISA-E Studien anhand dieser Standards ein Ländervergleich durchgeführt.

Die im Zusammenhang mit der Entwicklung und Prüfung von Bildungsstandards notwendigen Arbeitsschritte hatte das DESI-Projekt bei der Entwicklung, Erprobung und Skalierung seiner Tests bereits exemplarisch durchlaufen. Das IQB nutzt die Erfahrungen des DESI Projekts bei seiner weiteren Arbeit. Auf diese Weise wird das mit DESI begonnene Systemmonitoring für den Sprachunterricht als Daueraufgabe fortgeführt. Darüber hinaus ist ein indikatorenbasiertes Monitoring auf europäischer Ebene in Vorbereitung. Die Einführung der Bildungsstandards und die internationale Dimension des GERS haben Bewegung in das Nachdenken über die Zielsetzungen der Sprachvermittlung und den tatsächlichen Sprachenerwerb gebracht. Auf europäischer Ebene werden diese Themen sowohl im Europarat als auch in der Europäischen Union diskutiert. Der Europarat rief zu Projekten auf, in denen es um die soziale Kohäsion in Europa geht, wobei der kulturellen Kooperation „respecting the rich linguistic and cultural diversity of Europe“ eine zentrale Stellung zukommt. Unterricht, Training und Unterstützung für die Förderung interkulturell kommunikativer Sensibilität und Kompetenz im sprachlichen Bereich genießen derzeit auf der Agenda des Europarats eine hohe Priorität (Council of Europe 2003). Der GERS – ein Produkt auf Veranlassung des Europarats – dient unter anderem als Grundlage für die gegenseitige Anerkennung sprachlicher Qualifikationen. Für die EU hat der Spracherwerb eine besondere Bedeutung, weil Sprachkompetenzen hier auch als Voraussetzungen für wirtschaftlichen Austausch und Entwicklung betont werden. Dementsprechend hat die EU 2002 in Barcelona die Konstruktion eines Indikators für Sprachkompetenzen beschlossen, der verlässliche Daten über die Sprachenkenntnisse der jungen europäischen Generation liefern soll. Ein solcher Indikator soll bis 2011 erstellt werden und europäische Vergleiche von sprachlichen Kompetenzen ermöglichen. Ein wichtiger Vorläufer des europäischen Sprachenindikators ist das Projekt „Building a European Bank of Anchor Items for Foreign Language Skills“ (EBAFLS, DIPF 2007) in dem unter anderem Items aus den DESI Tests Hörverstehen Englisch und Leseverstehen Englisch verwendet wurden. Ziel dieses Projekts ist es standardisierte und im jeweiligen Land bereits erprobte Tests auf ihre europaweite Einsetzbarkeit zur Diagnostik von Sprachkompetenzen zu überprüfen und eine Itemdatenbank zu erstellen, die kulturfaire Sprachdiagnostik über Ländergrenzen hinweg ermöglicht. Dank dieser Entwicklungen kommt einem Projekt wie DESI nicht nur eine nationale Bedeutung zu, sondern es spielt auch im europäischen Kontext eine nicht unerhebliche Rolle.

1.4 Konsequenzen aus DESI für Schule und Unterricht

DESI zeigt, dass der Englischunterricht in Deutschland – vor allem in den Gymnasien – über eine sehr starke Leistungsspitze von 10 bis 15 Prozent der Schülerinnen und Schüler verfügt, deren Kompetenzen weit über das Anforderungsniveau der Lehrpläne hinausragen. Andererseits zeigt DESI vor allem in Hauptschulen, Integrierten Gesamtschulen und Schulen mit mehreren Bildungsgängen starke Defizite. Im Verlauf der neunten Jahrgangsstufe lässt sich ein deutlicher Kompetenzzuwachs nachweisen, wobei jedoch die Schere zwischen Gymnasiasten und Schülern anderer Bildungsgänge größer wird.

Im Fach Deutsch werden die Leistungsunterschiede zwischen den Bildungsgängen durch den Unterricht der neunten Jahrgangsstufe weder verstärkt noch verringert. Die Fähigkeit zum Umgang mit Grammatik und Sprachstil steigt in allen Schularten stark an, während für die Lesekompetenz und die Schreibfähigkeit nur schwache Zuwächse verzeichnet werden konnten. Auch am Ende des Schuljahres schreibt etwa die Hälfte der Schülerinnen und Schüler im Bildungsgang der Hauptschule und bei Integrierten Gesamtschulen Brieftexte, die ihren kommunikativen Zweck kaum erfüllen.

Aus diesen Befunden ergeben sich klare Herausforderungen für Schule und Unterricht. Besonders wichtig erscheint die Förderung von Jungen, die bei sprachlichen Kompetenzen – vor allem wenn Texte verfasst werden müssen – deutlich schwächere Leistungen erbringen als Mädchen. Um im Fach Deutsch den Leistungsrückstand von Jugendlichen mit nicht deutscher Erstsprache zu verringern, wären auch in der 9. Jahrgangsstufe zusätzliche Fördermaßnahmen angezeigt, nicht zuletzt im Blick auf den Wortschatz der Jugendlichen. Ermutigend ist, dass Schüler mit nicht deutscher Erstsprache – unter Berücksichtigung von Schulform und sozioökonomischem Hintergrund – bei der Rechtschreibung keine Benachteiligung zeigen und im Fach Englisch sogar einen Leistungsvorsprung haben.

Wie können die Herausforderungen gemeistert werden? Hierzu gibt DESI vielerlei Hinweise – auch solche, die weit verbreitete pädagogische Anschauungen in Frage stellen. So hat im Fach Deutsch die bloße Vielfalt an verwendeten Methoden und Textsorten keinen Zusammenhang mit dem Kompetenzzuwachs im Verlauf der neunten Jahrgangsstufe. Einzig die Auseinandersetzung mit Prosatexten wirkt sich spezifisch auf die Lesekompetenz aus. Bedeutsam hingegen ist, ob die Schülerinnen und Schüler den Eindruck haben, dass ihrer Lehrerin bzw. ihrem Lehrer sprachliche Kompetenzen wichtig sind. Klassen, deren Schüler berichten, Rechtschreibung, grammatisch richtiges Schreiben und andere sprachbezogene Fähigkeiten seien im Unterricht sehr wichtig, haben schon zu Beginn des Schuljahres im DESI-Test bessere Ergebnisse und steigern ihre Kompetenzen – auch im Lesen – im Verlauf des Jahres stärker als andere Klassen. Erfolgreicher Deutschunterricht zeichnet sich also auch in der Jahrgangsstufe neun durch klare Anforderungen im sprachlichen Bereich aus. Allgemeine Unterrichtsqualitätsmerkmale wie die effiziente Klassenführung (im Sinne der Vermeidung von Disziplinproblemen), die Einbeziehung der Schülerinnen und Schüler in den Deutschunterricht (Schülerorientierung) und ein kognitiv ak-

tivierender Umgang mit Hausaufgaben tragen zu höheren Kompetenzen in der Sprachbewusstheit bei.

Im Fach Englisch zeigt die DESI-Studie eine Reihe von Merkmalen auf, die der Unterricht haben sollte, um die kommunikative Kompetenz der Schülerinnen und Schüler optimal zu fördern:

- Eine hohe Wertigkeit der Korrektheit beim Sprechen, bezogen auf Aussprache und Grammatik, aber zugleich
- eine hohe Wertigkeit der Flüssigkeit im Sprachgebrauch,
- die Vermeidung der deutschen Sprache im Unterrichtsgespräch (während sie in anderen Situationen, z.B. bei der Instruktion für eine Gruppenarbeit, eher von Vorteil sein kann!),
- ein hoher Sprechanteil von Schülerinnen und Schülern,
- geduldigeres Warten auf Schülerantworten (über 3 Sekunden hinaus),
- Gelegenheiten für die Selbstkorrektur von Fehlern.

Auch im Fach Englisch kommt es also auf hohe sprachliche Anforderungen und auf die Qualität der Interaktion im Unterricht an, nicht auf die bloße Vielfalt von Methoden und Materialien. Darüber hinaus bestätigt DESI für das Fach Englisch allgemeine Faktoren der Unterrichtsqualität, die sich bereits in anderen Fächern als wirksam erwiesen haben:

- die Integration der Hausaufgaben in den Unterricht,
- eine effiziente Klassenführung,
- die Nutzung von Fehlern als Lerngelegenheiten,
- ausgeprägte Verständlichkeit der Lehreräußerungen,
- ein positives Unterrichtsklima,
- eine hohe Motivierungsqualität.

Auf der Ebene der Schule kann eine solche Entwicklung unterstützt werden, indem die Fachkollegen klare, gemeinsame Zielvorstellungen entwickeln. Das geteilte Ziel- und Aufgabenverständnis kann als Ansatzpunkt einer konsistenten Unterrichtsentwicklung und einer nachhaltigen Kompetenzförderung auf Schulebene genutzt werden; DESI belegt die Wirksamkeit solcher kohärenter Ziele. Auch konsequente Elternarbeit zeigt Erfolge. Eine wesentliche Verbesserung der Sprachkompetenz im Englischen kann zudem erreicht werden, wenn die Schule Fachunterricht, etwa für Geografie oder Biologie, in der Fremdsprache anbietet (sog. bilingualer Sachfachunterricht).

Schulen werden solche Aufgaben sicherlich nicht im Alleingang meistern können. Die Verbesserung der Lehreraus- und weiterbildung steht bereits auf der Agenda der Bildungspolitik. DESI zeigt hier, dass die Probleme des Fachs Englisch im Bildungsgang Hauptschule besonders dringlich sind. In der Mathematik und den Naturwissenschaften hat das Modellprogramm „SINUS“ vorgeführt, wie konzentrierte Maßnahmen zur Verbesserung von Schul- und Unterrichtsqualität schließlich auch in die Steigerung der Schülerkompetenzen münden. Im Bereich der Sprachen haben sich spezielle Fördermaßnahmen wie Sommercamps und der bilinguale

Sachfachunterricht auch empirisch bewährt. Diese Anstrengungen verdienen eine Verbreitung und Vertiefung.

1.5 Danksagung

Das Projekt DESI war nur realisierbar mit starker Unterstützung des von der KMK einberufenen wissenschaftlichen Beirats sowie vieler Fachleute in Ministerien und Landesinstituten sowie im Sekretariat der Kultusministerkonferenz. Auch mit dem Institut für Qualitätsentwicklung im Bildungswesen der KMK hat sich seit dessen Gründung eine enge Kooperation entwickelt.

Die Herausgeber dieses Bandes danken den Mitarbeiterinnen und Mitarbeitern der beteiligten Institute, die auch nach dem offiziellen Auslaufen des Projekts viel Arbeitszeit in dieses Publikationsvorhaben investiert haben. Insbesondere das Team am DIPF hat trotz starker Belastung in Anschlussprojekten im Jahr 2007 kontinuierlich an dem Buch gearbeitet. Neben Dr. Brigitte Steinert und Dr. Hermann Günter Hesse, die schon zu den Initiatoren der DESI-Studie gehörten, haben sich Dr. Johannes Hartig, Dipl.-Psych. Dominique Rauch und cand. psych. Svenja Vieluf in der Endredaktion des Bandes besondere Verdienste erworben.

Literatur

- Beck, B./Klieme, E. (2007): Sprachliche Kompetenzen – Konzepte und Messung. Weinheim: Beltz.
- Beck, B./Dahl, D. (2006): Sprachliche Kompetenzen von Schülerinnen und Schülern der neunten Jahrgangsstufe in Deutsch. Zentrale Befunde der Studie Deutsch-Schülerleistungen-International in Südtirol. [WWW-document]. URL http://www.schule.suedtirol.it/pi/downloads/desi_bericht_suedtirol.pdf. Zugriff: 20.12.07.
- Council of Europe (2003): Call for Proposals 2004-2007. Languages for social cohesion: Language education in a multilingual and multicultural Europe.
- DIPF (2007): The EBAFLS (European Bank of Anchor Items for Foreign Language Skills)-Project. [WWW-document]. URL http://www.dipf.de/bildungsforschung/biqua_ebafls.htm. Zugriff: 20.12.2007.
- Goethe Institut Inter Nationes, Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland (KMK), Schweizerische Konferenz der Kantonalen Erziehungsdirektoren & Österreichisches Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK) (2001): Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Berlin: Langenscheidt.
- Klieme, E. (2002): Was ist guter Unterricht? Ergebnisse der TIMSS-Videostudie im Fach Mathematik. In: Bergsdorf, W./Court, J./Eckert, M./Hoffmeister, H. (Hrsg.): Herausforderungen der Bildungsgesellschaft. 4. Ringvorlesung der Universität Erfurt. Weimar: Rhino Verlag, S. 89-113.
- Kultusministerkonferenz (2005): Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung. Neuwied: Luchterhand.
- Rupp, A./Vock, M./Harsch, C./Köller, O. (in Druck): The Development, Calibration, and Validation of Standards-based Tests for English as a First Foreign Language. Münster: Waxmann.

Bärbel Beck / Svenja Bundt / Jens Gommelka

2 Ziele und Anlage der DESI-Studie

Die Studie „Deutsch-Englisch-Schülerleistungen-International“ (DESI) untersucht die sprachlichen Leistungen von Neuntklässlern und die Unterrichtswirklichkeit in den Fächern Deutsch und Englisch. Etwa 11000 Schülerinnen und Schüler der neunten Jahrgangsstufe aller Schularten im allgemeinbildenden Schulwesen (ohne Sonderschule) wurden zu Beginn und am Ende des Schuljahres 2003/04 befragt und getestet; hinzu kamen Befragungen von Lehrkräften, Eltern und Schulleitungen sowie Videoaufnahmen im Englischunterricht. Als bundesweit repräsentative Untersuchung und durch ihr komplexes Design ermöglicht die Studie differenzierte Aussagen über Lehr-Lern-Bedingungen und den Erwerb sprachlicher Kompetenzen, die für Unterrichtspraxis und Bildungspolitik gleichermaßen wichtig sind.

2.1 Ziele

Grundlage der Schulleistungsstudie „Deutsch-Englisch-Schülerleistungen-International“ (DESI) war eine Ausschreibung der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) im März 1999. DESI ist die erste Schulleistungsstudie in der Verantwortung der KMK. Sie ist als pädagogische Qualitätsuntersuchung mit Schwerpunkt auf dem Unterricht zu charakterisieren und stellt keine Vergleichsstudie dar. Das Projekt wurde im Jahre 2001 an ein interdisziplinär zusammengesetztes Konsortium aus Fachdidaktikern, Psychologen und Schulforschern unter der Federführung des Deutschen Instituts für Internationale Pädagogische Forschung in Frankfurt am Main (DIPF) vergeben. Dem Konsortium gehören Wissenschaftler der Universitäten Augsburg, Berlin (Humboldt-Universität), Dortmund, Hamburg, Koblenz-Landau, Oldenburg und Osnabrück sowie des DIPF an.

Mit der Vorbereitung und Durchführung der Datenerhebung beauftragte das DESI-Konsortium das Data Processing Center der International Association for the Evaluation of Educational Achievement (IEA DPC) in Hamburg, das in Zusammenarbeit mit dem DIPF für eine an internationalen Standards orientierte Testdurchführung verantwortlich zeichnet.

Die Kultusministerkonferenz wollte mit dem Projekt Basisinformationen über den Kompetenzstand der Schülerinnen und Schüler der neunten Jahrgangsstufe in deutschen Schulen hinsichtlich des Englischen sowie hinsichtlich der aktiven Beherrschung des Deutschen gewinnen. Damit sollten die Ergebnisse der TIMS- und PISA-Studie, die sich auf Mathematik, Naturwissenschaften und die Lesekompetenz konzentrieren, ergänzt werden. Gegenstand der Untersuchung sind sowohl die reproduktiven als auch die produktiven, schriftlichen und (soweit in einer großen Studie

realisierbar) mündlichen Leistungen von Neuntklässlern in den Fächern Deutsch und Englisch als erster Fremdsprache.

Über die Modellierung von Kompetenzen und die Beschreibung des erreichten Kompetenzniveaus hinaus verbanden die Auftragnehmer mit DESI von Anfang an analytische Fragestellungen. Die Daten zum Kompetenzerwerb sollten auf der Grundlage von Rahmenmodellen der Schulwirkungsforschung mit einer Vielzahl von Faktoren auf schulischer, klassenspezifischer und individueller Ebene verknüpft werden. Dadurch sollte DESI Erklärungsansätze für unterschiedliche Kompetenzniveaus sowie Grundlagenwissen für bildungspolitische Aktivitäten, für die Revision von Curricula, für die Gestaltung von Lehrtexten und Unterrichtsmaterialien, für die Lehrerbildung und -fortbildung und vor allem für die Unterrichtsgestaltung liefern.

Derartige Fragestellungen lassen sich in querschnittlich angelegten Studien nicht beantworten, weil Kompetenzen aus langjährigen Prozessen resultieren, deren Ausgangs- und Rahmenbedingungen sich im Nachhinein höchst unzureichend rekonstruieren lassen. Effekte der schulischen und häuslichen Lernumwelt können nur in längsschnittlichen Untersuchungen mit mindestens zwei Messzeitpunkten abgegrenzt werden. Dementsprechend war DESI als Längsschnittstudie mit einer differenzierten Erfassung des Unterrichts und der Voraussetzungen im Elternhaus über ein Jahr hinweg konzipiert.

Während DESI konzipiert und umgesetzt wurde, begann die KMK mit der Arbeit an länderübergreifenden Bildungsstandards. Die Studie gewinnt somit zusätzliche Bedeutung als Vorarbeit zur Konkretisierung und empirischen Überprüfung von Bildungsstandards in diesen beiden Fächern.

2.2 Untersuchungspopulation und Ziehung der Stichprobe

Zielpopulation

Die Zielpopulation der DESI-Studie stellen alle Schülerinnen und Schüler innerhalb Deutschlands dar, die sich zum Zeitpunkt der Untersuchung in der neunten Jahrgangsstufe an allgemeinbildenden Schulen befanden¹. Förderschüler wurden aufgrund der inhaltlichen Ausrichtung der Studie nicht berücksichtigt.

Stichprobenplan

Um den Stichprobenplan aufzustellen, wurde die Anzahl der zu ziehenden Schulen proportional zur Anzahl der Neuntklässler in den Bundesländern auf die einzelnen Bundesländer verteilt, so dass kleine Bundesländer (wie z.B. die Stadtstaaten) nur

1 Die Schüler stammten aus den Schularten Hauptschule, Realschule, Integrierte Gesamtschule, Schulen mit mehreren Bildungsgängen (=MBG; wie Sekundar-, Mittel- oder Regelschule) und Gymnasium.

mit wenigen Schulen in der Stichprobe vertreten sind, große Länder dagegen mit mehr Schulen (explizite Stratifizierung).

Die Anzahl der Schulen pro Schulart innerhalb eines Bundeslandes ergibt sich näherungsweise proportional aus dem Anteil der Schülerinnen und Schüler innerhalb der jeweiligen Schulart in einem Bundesland (implizite Stratifizierung). Dabei ist zu beachten, dass in kleinen Ländern (z.B. Bremen), die nur mit wenigen Schulen in der Stichprobe vertreten sind, die Zahl der zu ziehenden Schulen unter der Anzahl möglicher Schularten lag, so dass nicht jede Schulart repräsentiert wird, obwohl jede Schule auch in diesen Bundesländern die gleiche Chance hatte, für die Untersuchung ausgewählt zu werden.

Einen zusätzlichen Untersuchungsaspekt stellt in DESI die Gruppe der bilingual unterrichteten Schülerinnen und Schüler dar. Um eine Mindeststichprobengröße für diese Personengruppe zu erzielen, gingen Schulen mit bilingual unterrichteten Schülern überproportional in die Stichprobe ein (*oversampling*). Beim *oversampling* der bilingualen Schulen wurde darauf geachtet, dass sich die Anzahl der auszuwählenden bilingualen Schulen ungefähr proportional zum Anteil der Schulen mit bilingualem Unterrichtsangebot im jeweiligen Bundesland bestimmt. Generell gibt es bilingualen Unterricht in Deutschland in erster Linie an Gymnasien, so dass vor allem diese Schulart in der DESI-Stichprobe der bilingualen Schulen vertreten ist. Tabelle 2.1 zeigt die Verteilung der pro Bundesland gezogenen Schulen nach Schulart².

Stichprobenziehung

Im ersten Schritt (Schulstichprobenziehung) wurden entsprechend dem DESI-Stichprobenplan per Zufallsziehung 180 Schulen ausgewählt. Zusätzlich wurden 40 bilinguale Schulen gezogen. In einem zweiten Schritt (Klassenstichprobenziehung) wurden – ebenfalls in einem Zufallsverfahren – aus allen Klassen der neunten Jahrgangsstufe innerhalb einer gezogenen Schule zwei komplette Klassen für die Teilnahme an DESI ausgewählt. In Integrierten Gesamtschulen, an Schulen mit mehreren Bildungsgängen und in den Hauptschulen einzelner Bundesländer, in denen das Fach Englisch leistungsdifferenziert im Kurssystem unterrichtet wird, wurden anstelle von zwei kompletten neunten Klassen in der Hälfte der Schulen zwei Englischkurse mit den entsprechenden Schülerinnen und Schülern als Testgruppe ausgewählt. In der anderen Hälfte dieser Schulen wurden – wie in den übrigen

2 GY-Gymnasium, HS-Hauptschule, IGS-Integrierte Gesamtschule, MBG-Schule mit mehreren Bildungsgängen, RS-Realschule: Diese Abkürzungen gelten für den gesamten Sammelband.

3 An diesen Schulen gibt es in Englisch leistungsdifferenzierte Grund- und Erweiterungskurse oder auch A-, B-, C-Kurse.

Schulen – zwei komplette neunte Klassen für den Test gezogen, in denen das Fach Deutsch im Klassenverband unterrichtet wurde.⁴

Durch dieses Zufallsverfahren befinden sich in der Stichprobe auch einige Schülerinnen und Schüler – überwiegend an Gymnasien –, die Englischunterricht erst nach der Jahrgangsstufe 5 aufgenommen haben, d.h. zumeist in Klasse 7. Diese Gruppe ist mit 3,8% der Gesamtstichprobe jedoch sehr klein.

Tabelle 2.1: DESI-Schulstichprobenplan: Anzahl der Schulen nach Bundesländern und Schulformen.

Bundesland	GY	HS	IGS	MBG	RS	Gesamt
Brandenburg	2		4		1	7
Berlin	3	1	2		1	7
Baden-Württemberg	8	7	1		7	23
Bayern	9	9			9	27
Bremen		1			1	2
Hessen	6	2	2		5	15
Hamburg	2	1		1		4
Mecklenburg Vorpommern	1			3	1	5
Niedersachsen	13	6	1		7	27
Nordrhein-Westfalen	21	12	6		17	56
Rheinland-Pfalz	5	3		1	3	12
Schleswig-Holstein	2	2	1		3	8
Saarland			1	1		2
Sachsen	4			7		11
Sachsen-Anhalt	2			5		7
Thüringen	3			4		7
Gesamt	81	44	18	22	55	220
davon bilingual	33				7	40

Klassen und Schüler im Längsschnitt

Die zwei pro Schule gezogenen Klassen bzw. Kurse wurden zu zwei Messzeitpunkten – zu Beginn und am Ende des Schuljahres 2003/04 – untersucht. Der Längsschnitt umfasste damit in der Regel dieselben Schülerinnen und Schüler. Allerdings fielen einzelne Schüler, die im Laufe des Jahres in eine nicht an DESI beteiligte Parallelklasse

4 Diese Besonderheit der Deutsch-Klassenziehung bzw. Englisch-Kursziehung führte auch dazu, dass in diesen Schulen jeweils nur die Deutsch- oder die Englischlehrkräfte an der Lehrerbefragung teilnehmen konnten und nicht – wie in den übrigen Schulen – sowohl die Deutsch- als auch die Englischlehrkraft der jeweils gezogenen Klasse, da die Bezugsgruppe der Schüler an den Schulen mit Kurssystem für beide Fächer/Fachlehrkräfte nicht die gleiche gewesen wäre.

gewechselt waren, zum zweiten Messzeitpunkt aus der Erhebung heraus (483 Schülerinnen und Schüler). Kam ein Schüler im Laufe des Schuljahres neu in eine DESI-Klasse hinzu, so wurde er zum zweiten Messzeitpunkt in die Untersuchung einbezogen (insgesamt belief sich die Anzahl dieser Fälle auf 497).

Auswahl der Klassen in den bilingualen Schulen

In der Regel wird bilingualer Sachfachunterricht an den als bilingual bezeichneten Schulen nur in einem Teil der Parallelklassen der neunten Jahrgangsstufe erteilt. Um den Anteil der bilingual unterrichteten Schülerinnen und Schüler in der Stichprobe zu maximieren, wurden bei der Ziehung der Klassen in den bilingualen Schulen diejenigen Klassen, in denen alle oder zumindest ein Teil der Schüler bilingual unterrichtet werden, mit Sicherheit in die Erhebung aufgenommen, sofern weniger als drei Klassen mit bilingual unterrichteten Schülern an der Schule vorhanden waren. Gab es nur eine Klasse mit bilinguaem Unterricht, wurde aus den übrigen nicht bilingualen neunten Klassen der Schule eine weitere Klasse ausgewählt und in die Untersuchung einbezogen.

Auswahl der Schülerinnen und Schüler für den SET-10-Test

Für den Test zur Sprachproduktion im Englischen (Ordinate Spoken English Test SET-10, vgl. Kapitel 16) wurden aus jeder an DESI beteiligten Klasse bzw. aus jedem Kurs – ebenfalls in einem Zufallsverfahren – jeweils drei Mädchen und drei Jungen ausgewählt. Zusätzlich wurden pro Klasse je ein Junge und ein Mädchen als Ersatzkandidaten ausgewählt, um bei kurzfristiger Erkrankung eines ehemals vorgesehenen Schülers diesen zu ersetzen.

Ausschluss von Schülerinnen und Schülern

Analog zu anderen Schulleistungsstudien war es im Rahmen von DESI nicht zugelassen, Schüler nur aufgrund von geringer Leistungsfähigkeit vom Test auszuschließen. Allerdings gab es vergleichbar mit internationalen Standards aus PISA und IGLU auch in DESI klar definierte Ausschlusskriterien, nach denen folgende Personen vom Test ausgeschlossen werden konnten:

- Schüler mit einer permanenten körperlichen Behinderung, die ihnen die Teilnahme an der Testsituation unmöglich machte.
- Schüler, die emotional oder geistig nicht in der Lage waren, auch nur den allgemeinen Anweisungen des Tests zu folgen, oder die durch die Testsituation in unzumutbarem Maße emotional belastet worden wären.
- Schüler, deren Erstsprache nicht deutsch war, und die weniger als ein Jahr in deutscher Sprache unterrichtet worden waren.
- Innerhalb der DESI-Studie wurden 0.7% der Schülerinnen und Schüler aufgrund eines der oben genannten Kriterien von der Untersuchung ausgeschlossen.

Gewichtung

Mit der Gewichtung wird die Tatsache ausgeglichen, dass durch den stratifizierten Stichprobenplan in DESI nicht alle Schülerinnen und Schüler der Jahrgangsstufe 9 in Deutschland mit der gleichen Wahrscheinlichkeit in die Stichprobe gelangten. Um einzelne Ziehungsgruppen (z.B. bilingual unterrichtete Schüler) bei den Analysen nicht über- oder unterrepräsentiert eingehen zu lassen, erhielten sie ein entsprechendes Gewicht, das sowohl die verschiedenen Ebenen der Stichprobenziehung (Schulen, Klassen, Schüler) als auch die unterschiedlichen Beteiligungsdaten auf diesen Ebenen mit einbezieht. In der DESI-Studie war es sehr wichtig, mit Gewichten zu rechnen, um das oversampling der bilingualen Schulen auszugleichen. Der Ausgleich für unterschiedlich hohe Beteiligungsdaten wird auch als Adjustierung bezeichnet.

Im ersten Schritt fand eine Gewichtung nach explizitem Stratum (d.h. nach Bundesland) statt. Danach erfolgte eine Adjustierung auf Schulebene, indem die Nichtteilnahme einzelner Schulen miteinbezogen wurde. Dieses kam in DESI nur in einem Fall vor (vgl. Abschnitt zu den Beteiligungsdaten). Für die Klassengewichtung wurden im nächsten Schritt die Auswahlwahrscheinlichkeiten der beiden neunten Klassen bzw. Kurse und die Beteiligungsrate in diesen Klassen/ Kursen verrechnet. In DESI mussten hierbei unbedingt bilinguale und nicht bilinguale Schulen unterschieden werden. Für die Gewichtung von Schülern innerhalb von Klassen betrug die Auswahlwahrscheinlichkeit immer 1, da komplette Klassen oder Kurse gezogen wurden. Es fand jedoch auch auf der Schülerebene eine Adjustierung der Gewichte statt, bei der die Nichtteilnahme auf dieser Ebene berücksichtigt wurde. Die Güte der bei DESI vorgenommenen Gewichtung zeigt sich daran, dass die Populationsverhältnisse durch die Gewichte entsprechend internationalen Standards sehr gut abgebildet wurden.⁵ Die Befunde des vorliegenden Bands wurden – mit Ausnahme von Kapitel 35, das explizit den bilingualen Unterricht analysiert – mit diesen Gewichtungen errechnet und bilden somit die Verhältnisse in der Zielpopulation (Neunte Jahrgangsstufe an allgemeinbildenden Schulen in Deutschland) repräsentativ ab.

Gewichtung der SET-10-Test-Schüler

Für diejenigen Schülerinnen und Schüler, die am SET-10-Test teilgenommen haben, war die Auswahlwahrscheinlichkeit nicht gleich 1, da innerhalb einer Klasse eine Substichprobe gezogen wurde. Da für den SET-10-Test jeweils drei Mädchen und drei Jungen ausgewählt wurden, richtet sich der Schülerfaktor in diesem Fall nach der Klassengröße: Je größer die Klasse ist, desto größer ist auch der Schülerfaktor des jeweiligen Schülers.

5 Geschätzte Population: 916430 Schülerinnen und Schüler; offizielle Schülerzahl laut statistischem Bundesamt, Fachserie 11, Reihe 1, Bildung und Kultur, Allgemein bildende Schulen, Schuljahr 2002/2003: 934397 Schülerinnen und Schüler; Abweichung: 17967 Schülerinnen und Schüler).

Gewichtung und Längsschnitt

Angesichts der Tatsache, dass in der DESI-Längsschnittuntersuchung einzelne Testmodule nur zu einem Messzeitpunkt eingesetzt wurden und es weiterhin eine bestimmte Anzahl von Schülern gab, die nur zu einem Messzeitpunkt am Test teilgenommen haben, stellte sich die Frage, wann ein Schüler als ‚teilgenommen‘ gewertet wird. Hierzu wurde für DESI festgelegt, dass ein Schüler im Rahmen der DESI-Längsschnittauswertung als ‚Teilnehmer‘ gilt, wenn er für mindestens ein Testmodul auswertbare Leistungen erzielt hat. Diese Vorgehensweise ermöglicht es, alle Analysen auf Schülerebene mit nur einer Gewichtungsvariable durchzuführen. Fehlende Werte bezüglich einzelner Testmodule wurden dabei mit Hilfe der multiplen Imputation (vgl. Kapitel 3) ersetzt.

2.3 Testentwicklung und Design der Untersuchung⁶

Für die Umsetzung der mit der DESI-Studie verbundenen Ziele war die Entwicklung lehrplanorientierter Tests zur Erfassung sprachlicher Kompetenzen im Deutschen und Englischen notwendig. Die Testentwicklung beinhaltet im Wesentlichen die nachfolgenden Schritte.

Schritt 1: Theoretisches Rahmenkonzept und Präpilotierung

Der erste Schritt diente der Definition des Konstrukts, der sprachtheoretischen, didaktischen und curricularen Einordnung der Leistungsmodule, der Recherche nach vorhandenen und/oder der Entwicklung eigener Operationalisierungsansätze und schließlich der informellen Erprobung prototypischer Aufgaben. Da in DESI der Schwerpunkt der Analysen auf dem Unterricht in den Fächern Deutsch und Englisch lag, wurden die DESI-Leistungstests an den Inhalten der Lehrpläne für die neunte Jahrgangsstufe der Sekundarstufe I ausgerichtet, also kriteriumsorientiert konstruiert. Zu diesem Zwecke wurden detaillierte Lehrplananalysen durchgeführt, die das Ziel hatten, länder- und bildungsgangübergreifende Inhalte zu identifizieren. Diese Lehrplaninhalte stellen institutionelle Normen für die Entwicklung und Bewertung sprachlicher Kompetenzen dar. Die DESI-Ergebnisse erlauben eine Antwort auf die Frage, in welchem Grade die Neuntklässler diesen curricularen Normen nahe kommen. Im Sinne der Testökonomie wurden über alle Bildungsgänge hinweg im Wesentlichen die gleichen Aufgaben gestellt.⁷ Das DESI-Kompetenzkonzept stellt durch den direkten curricularen Bezug in Deutschland eine Innovation auf dem

6 Die Fragebogenentwicklung wird in den Kapiteln zu den Bedingungsfaktoren sprachlicher Kompetenzen von den dortigen Autoren dargestellt.

7 Eine Ausnahme bilden die Englischtests Hörverstehen, Textrekonstruktion und Leseverstehen. In diesen Tests waren neben übergreifenden auch bildungsgangspezifische Aufgaben enthalten, die sich in der Schwierigkeit der Anforderungen unterscheiden, aber auf einer gemeinsamen Kompetenzskala abgebildet werden können.

Gebiet der Messung von Schülerkompetenzen dar und unterscheidet sich darin vom Literacy-Konzept der PISA-Studien.

Schritt 2: Testkonzeption und Pilotierung

Die informelle Erprobung erster prototypischer Aufgaben führte schließlich zu solchen Testaufgaben, die in die Pilotierung Eingang fanden. Dieser Schritt schloss schließlich Überlegungen zum Auswertungskonzept, zum Pilotierungsdesign und zur Durchführung der Pilotierung ein. Die Pilotierung fand in der Zeit von September bis Oktober 2002 in Kooperation mit dem IEA DPC statt (Klieme u.a. 2003) und mündete in die Konstruktion der Erhebungsinstrumente und die Erstellung der Druckvorlagen für die Hauptuntersuchung.

Schritt 3: Hauptuntersuchung: Durchführung (vgl. Punkt 2.4.)

Die DESI-Untersuchung wurde wesentlich durch ihr Längsschnittsdesign charakterisiert. Der erste Messzeitpunkt lag am Anfang der neunten Jahrgangsstufe (September/Oktober 2003); hier fand ein Untersuchungstag statt. Der zweite Messzeitpunkt lag am Ende der neunten Jahrgangsstufe (Mai bis Juli 2004). Die Erhebungen erstreckten sich nun über zwei aufeinander folgende Testtage.⁸ Zwischen den beiden Messzeitpunkten wurde in 105 Klassen die Videostudie des Englischunterrichts durchgeführt.

Die Vorgabe der Leistungstests erfolgte in DESI in einem Matrix-Design, bei dem jeder Schüler nur einen Teil der Aufgaben jedes Tests bearbeitet. Dieses in *large scale* Studien übliche Vorgehen hat den untersuchungsökonomischen Vorteil, dass insgesamt eine breite Menge von Testaufgaben eingesetzt werden kann und die individuelle Testzeit dennoch begrenzt bleibt. Für DESI ist das Matrix-Design aus einem zweiten Grund unerlässlich: Sprachkompetenzen wurden zu Beginn und am Ende der neunten Jahrgangsstufe getestet, um den Zuwachs innerhalb der neunten Jahrgangsstufe ermitteln zu können. Hierbei können dieselben Schüler zu zwei Zeitpunkten nicht mit denselben Aufgaben getestet werden, da sie sich an diese noch erinnern könnten. Die Vorgabe der Tests im Matrix-Design erlaubt es, jedem Schüler zu beiden Zeitpunkten verschiedene Aufgaben vorzugeben.

Die Testhefte wurden so konzipiert, dass pro Klasse und in den meisten Fällen pro Leistungsmodul vier Testheftvarianten mit jeweils überlappenden Testaufgaben existierten. Für einige Englischtests wurden darüber hinaus bildungsgangspezifisch konzipierte Testheftversionen zusammengestellt, um ein breites Leistungsspektrum adäquat abdecken zu können. Die Darstellung des allgemeinen Prinzips des rotier-

8 Für jeweils sechs Schüler einer jeden DESI-Klasse fand ein zusätzlicher dritter Untersuchungstag statt, der dem Einsatz des computergestützten SET-10-Tests zur Erfassung der mündlichen Sprachkompetenz im Englischen diente. Dieser dritte Testtag schloss sich aber nicht direkt an die beiden regulären Testtage an, sondern wurde aus organisatorischen Gründen ca. eine Woche später durchgeführt.

ten Testdesigns pro Klasse für den ersten und zweiten Messzeitpunkt findet sich in Tabelle 2.2.

Tabelle 2.2: Schema für das rotierte Testdesign der Hauptuntersuchung pro Klasse und Leistungsmodul.

Zeitpunkt	Teilstichproben pro Klasse	Testteile pro Klasse			
		A	B	C	D
Anfang des Schuljahres	1	A	B		
	2		B	C	
	3			C	D
	4	A			D
Ende des Schuljahres	1			C	D
	2	A			D
	3	A	B		
	4		B	C	

Mit dem in Tabelle 2.2 veranschaulichten Design wurde gewährleistet, dass

- jeder Schüler zu beiden Zeitpunkten unterschiedliche Aufgaben bearbeitete,
- die Items innerhalb jedes der beiden Zeitpunkte verankert waren, d.h. eine Prüfung auf Homogenität aller Aufgaben innerhalb jedes Zeitpunkts möglich war,
- alle Schüler etwa gleich viele Aufgaben zu beiden Zeitpunkten bearbeiteten,
- alle Aufgaben zu beiden Zeitpunkten vorgegeben wurden, damit auch eine Prüfung auf Gleichheit der Parameter zwischen den Zeitpunkten möglich war,
- jeder Testteil mit jedem kombiniert wurde.

Die Schätzung vergleichbarer Kompetenzwerte für verschiedene Testhefte und verschiedene Zeitpunkte ist mit Methoden der Item-Response-Theorie (IRT) möglich, vergleiche hierzu Kapitel 3.

Für alle Schulen umfassten die Testhefte jeweils Testteile aus den Bereichen Deutsch und Englisch und schlossen mit der Bearbeitung eines Schülerfragebogens ab. Darüber hinaus wurden zum zweiten Messzeitpunkt in allen Klassen die drei folgenden Tests eingesetzt:

- Metakognitives Wissen über Textverarbeitung (Schneider/Schlagmüller 2001),
- Kognitiver Fähigkeitstest, Untertests Figuren- und Wortanalogien (Heller/ Perleth 2000).

Die Tabellen 2.3 und 2.4 geben einen Überblick über die eingesetzten Tests und Fragebögen.

Tabelle 2.3: DESI-Tests zu den zwei DESI-Messzeitpunkten.

Instrumente	1. Messzeitpunkt	2. Messzeitpunkt	Varianten	Bemerkung
Tests Deutsch	Sprachbewusstheit	Sprachbewusstheit	4	
	Leseverstehen	Leseverstehen	4	
	Schreiben	Schreiben	4	
		Argumentation	4	
		Wortschatz	4	
		Rechtschreiben	1	
Tests Englisch	Sprachbewusstheit- Soziopragmatik	Sprachbewusstheit- Grammatik	1	
	Hörverstehen	Hörverstehen	8	
	Textrekonstruktion: C-Test	Textrekonstruktion: C-Test	12	
		Leseverstehen	6	
		Schreiben	2	
		Sprechen: Spoken- Englisch-Test-10	1	6 Schüler pro Klasse
		Interkulturelle Kompetenz	1	
weitere Tests		Kognitiver- Fähigkeitstest: Untertests Wortanalogien und Figurenanalogien	1	
		Lesegeschwindigkeit	1	
		Metakognitives Wissen über Textverarbeitung	1	

Durch Kombination und Optimierung aller Bedingungen ergaben sich für den ersten Messzeitpunkt 20 und für den zweiten Messzeitpunkt 27 verschiedene Testhefte.

Die Abfolge der eingesetzten Tests wurde aus testökonomischen Gründen konstant gehalten. Sowohl zum ersten als auch zum zweiten Messzeitpunkt waren zuerst die Englisch- und dann die Deutschtets zu bearbeiten. Da es sich bei allen DESI-Tests um Niveautests handelt und die Beobachtungen der Testleiter übereinstimmend das Ergebnis erbrachten, dass durch die Zeitregie keinerlei Bearbeitungsdruck auf die Schüler ausgeübt wurde, kann man davon ausgehen, dass sich die Abfolge der Tests nicht nachteilig auf die Testleistung ausgewirkt hat. Diese Aussage wird durch die testbezogene Selbsteinschätzung der Schüler bestätigt.

Tabelle 2.4: DESI-Fragebögen zu den zwei DESI-Messzeitpunkten.

Instrumente	1. Messzeitpunkt	2. Messzeitpunkt	Varianten	Bemerkung
	Schülerfragebogen Selbsteinschätzung	Schülerfragebogen Selbsteinschätzung	1	nach jedem DESI-Testteil eine Frage
Fragebögen Schüler	Schülerfragebogen kurz Lernen Deutsch/ Englisch	Schülerfragebogen lang Lernen Deutsch/ Englisch	2	in einer Hälfte der Klasse Deutsch; in der anderen Englisch
	Schülerfragebogen Einstellung zu Sprache		1	nicht in Hauptschulen
		Schülerfragebogen Unterrichtswahr- nehmung Deutsch/Englisch	2	in einer Hälfte der Klasse Deutsch; in der anderen Englisch
Fragebögen Lehrer	Lehrerfragebogen kurz Deutsch/Englisch	Lehrerfragebogen lang Deutsch/Englisch	2	
Fachkollegien		Fragebogen Fachkollegien Deutsch/Englisch	2	
Schulleitungen		Fragebogen Schulleitungen	1	
Eltern		Elternfragebogen	1	

Schritt 4: Hauptuntersuchung: Auswertung und Ergebnisdarstellung

Die differenzierte Auswertung der Daten wurde durch vor- und aufbereitende Arbeiten im IEA DPC ermöglicht. Sehr bewährt hat sich in diesem Zusammenhang das Erarbeiten und Erproben der Auswertungsroutinen anhand von Modelldatensätzen, die das IEA DPC vorab dem DIPF zur Verfügung gestellt hatte. Die Richtlinien für die Auswertungen wurden federführend vom DIPF erarbeitet, im Konsortium diskutiert und durch kontinuierliche bi- und multilaterale Abstimmungen umgesetzt. Ein Zentrum für die Auswertung der Leistungsdaten war dabei im DIPF; die Auswertung der Daten zu den beiden Schreibmodulen (Textproduktion Deutsch/Englisch) erfolgte an der Humboldt-Universität in Berlin. Die Fragebogendaten wurden von den verantwortlichen Autoren der Fragebogeninstrumente in Kooperation mit dem DIPF ausgewertet.

2.4 Durchführung der Erhebung

Mit der Vorbereitung und Durchführung der Datenerhebung hat das DESI-Konsortium das IEA DPC in Hamburg beauftragt, das in Zusammenarbeit mit dem DIPF für eine an internationalen Standards orientierte Testdurchführung verantwortlich zeichnete.

In die Vorbereitung der Untersuchung in den einzelnen Bundesländern wurden auch die jeweiligen Landesdatenschutzbeauftragten einbezogen. Um im Vorwege zu überprüfen, ob die Anonymität aller Befragten und die Datensicherheit gewährleistet sind, erhielten die Datenschutzbeauftragten relevante Unterlagen wie z.B. die Schüler- und Elternanschriften und die einzusetzenden Fragebögen. Erst nach der Begutachtung dieser Unterlagen wurde die Datenerhebung in den Bundesländern offiziell vom jeweiligen Kultusministerium genehmigt.

Vorbereitung

Nach der Ziehung der Schulstichprobe durch das IEA DPC stellten die DESI-Ansprechpartner in den Kultusministerien den Erstkontakt zu den für die Teilnahme bestimmten Schulen her. 219 der ausgewählten 220 Schulen erklärten sich zur Teilnahme an der DESI-Untersuchung bereit. Über Anschreiben, Manuale und Schulleiterinformationsveranstaltungen in den einzelnen Bundesländern wurden die Schulen über die Zielsetzungen der Studie und organisatorische Fragen der Erhebung informiert. Innerhalb jeder Schule bereitete ein Schulkoordinator/eine Schulkoordinatorin die Erhebung vor. So führte er/sie zur Vorbereitung der Klassenstichprobenziehung in der Klassenliste alle in Frage kommenden Klassen bzw. Kurse der Schule auf und sandte die Liste im Anschluss daran an das IEA DPC. Das IEA DPC nahm daraufhin die Klassenstichprobenziehung vor und teilte den Schulen mit, welche Klassen bzw. Kurse für die Untersuchung ausgewählt wurden. Alle Schülerinnen und Schüler der gezogenen Klassen erhielten über ihren Schulkoordinator ein von der Projektleitung am DIPF bereitgestelltes Schüler- und Elternansreiben und ein Informationsfaltblatt über die Studie. Auf diese Weise wurden die Schülerinnen und Schüler und ihre Eltern über den Test informiert und in allen Bundesländern, wo dieses notwendig war, wurde die schriftliche Einwilligung der Eltern zur Teilnahme ihrer Kinder eingeholt.

Test und Befragung der Schülerinnen und Schüler

Die Tests an den Schulen führten vom IEA DPC geschulte externe Testleiter nach einem standardisierten Testleiterskript durch, in dem alle Zeitvorgaben und Anweisungen an die Schüler genauestens niedergelegt waren. In Testsitzungsprotokollen konnten zudem besondere Vorkommnisse während der Testdurchführung protokolliert werden. In der Regel wurden zwei Testleiter pro Schule eingesetzt, so dass die Tests in den beteiligten Klassen parallel abliefen.

Nach Abschluss der Tests in den Schulen wurde das gesamte Testmaterial von den Testleitern an das IEA DPC geschickt, wo die weitere Datenerfassung stattfand.

Befragungen der Eltern, Lehrkräfte und Schulleitungen

Etwa zwei Wochen vor den Testterminen erhielten die Schulen vom IEA DPC die Fragebögen für Schulleitungen und Lehrkräfte bzw. die Fachkollegien Deutsch und Englisch. Die Fragebögen sollten bis zum Testtermin ausgefüllt werden und wurden von den externen Testleitern in Empfang genommen und zusammen mit dem übrigen Testmaterial an das IEA DPC weitergeleitet.

Die Elternfragebögen wurden den beteiligten Schülerinnen und Schülern am ersten Testtag des zweiten Messzeitpunktes ausgeteilt. In einem Anschreiben wurden die Eltern gebeten, den Fragebogen auszufüllen und ihn ihren Kindern möglichst umgehend wieder mit in die Schule zu geben. Dort wurden die zurückgekommenen Elternfragebögen vom Schulkordinator eingesammelt und ca. eine Woche nach dem Test zusammen mit nachträglich eingegangenen Lehrer- oder Schulfragebögen an das IEA DPC versandt.

Kodierung und Dateneingabe

Nach Eingang und Sichtung des bearbeiteten Testmaterials begann die Datenerfassung im IEA DPC. Dieses geschah auf zwei Wegen: Die in den einzelnen Testmodulen enthaltenen Aufgaben mit offenen Antwortformaten wurden zunächst kodiert, d.h. von im IEA DPC geschulten Kodierern nach vorgegebenen Kodieranweisungen bewertet.⁹ Die Ergebnisse der Kodierung wurden dabei auf Kodierblätter übertragen, die im Anschluss daran – wie auch alle Fragebögen und Testteile mit ausschließlich geschlossenen Antwortformaten (*multiple-choice*) – per Scannertechnik eingelesen wurden. Die Datensätze wurden um offensichtliche Erfassungsfehler bereinigt und so aufbereitet, dass sie an die Projektleitung am DIPF weitergegeben werden konnten, wo die Daten für weitere Analysen vorbereitet wurden.

Beteiligungsraten

Noch vor Beginn der konkreten Testvorbereitung wurden 14 der 220 laut Stichprobenplan gezogenen Schulen durch ebenfalls per Zufall ausgewählte Ersatzschulen ersetzt, da die Untersuchung in diesen Schulen zumeist aus schulorganisatorischen Gründen wie laufende Baumaßnahmen nicht stattfinden konnte. Insgesamt beteiligten sich 219 Schulen zu beiden Untersuchungszeitpunkten an der DESI-Studie. Eine Schule fiel aufgrund nur sehr weniger vorliegender Elterngenehmigungen zu einem Zeitpunkt aus der Stichprobe heraus, zu dem aus organisatorischen Gründen keine Ersatzschule mehr benannt werden konnte.

⁹ Die Kodierung von Deutsch- und Englischtextproduktion erfolgte in Eigenverantwortung der Testautoren in Berlin und Augsburg.

Von insgesamt 10639 Schülern zu Beginn der neunten Jahrgangsstufe beteiligten sich 94% an den Leistungstests und 92% an der Beantwortung der Fragebögen.¹⁰ Von 10632 Schülern am Ende der zweiten Jahrgangsstufe beteiligten sich 95% an den Leistungstests und 92% an der Beantwortung der Fragebögen. Dies ist gemessen an anderen Schulleistungsstudien eine sehr gute Teilnahmequote. Dieselbe Einschätzung trifft auch auf die Lehrer- und Schulleiterbefragungen zu.

Im Folgenden sind die Teilnahmequoten für Schüler und Eltern (Abbildung 2.1), Deutsch- und Englischlehrer (Abbildung 2.2) sowie Schulleiter und Fachkollegien (Abbildung 2.3) dokumentiert.

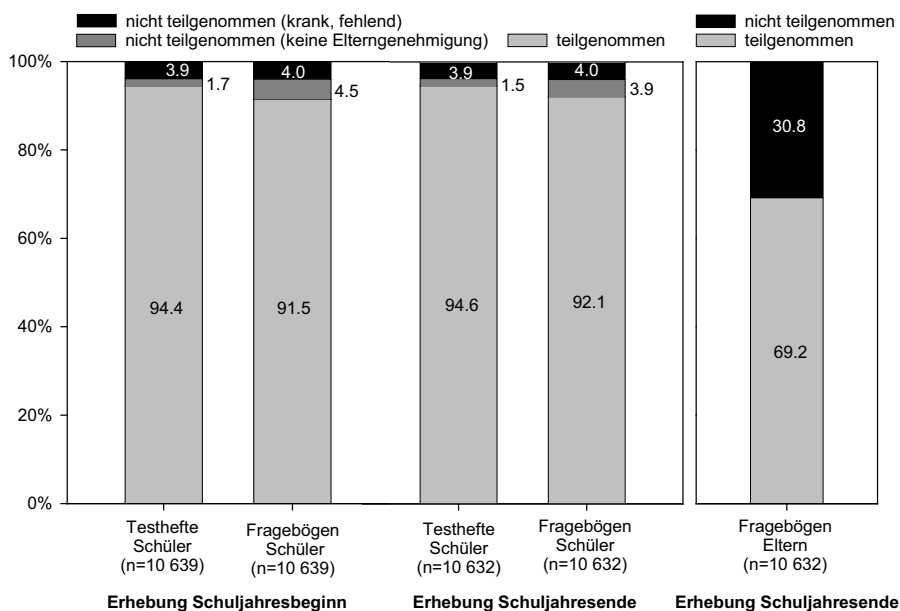


Abbildung 2.1: Teilnahmequoten für Schüler und Eltern.

¹⁰ Die etwas geringeren Teilnahmequoten bei der Bearbeitung der Schülerfragebögen gegenüber der Beteiligungsrate an den Leistungstests kommt dadurch zustande, dass die Bearbeitung der Leistungstests in der Mehrzahl der Bundesländer (Berlin, Brandenburg, Hamburg, Bremen, Hessen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen, Sachsen-Anhalt, Thüringen) mittlerweile verpflichtend für die Schülerinnen und Schüler ist, während für die Bearbeitung des Schülerfragebogens in den meisten Bundesländern eine schriftliche Einverständniserklärung der Eltern vorliegen muss (Rechtlicher Stand: Frühjahr 2004).

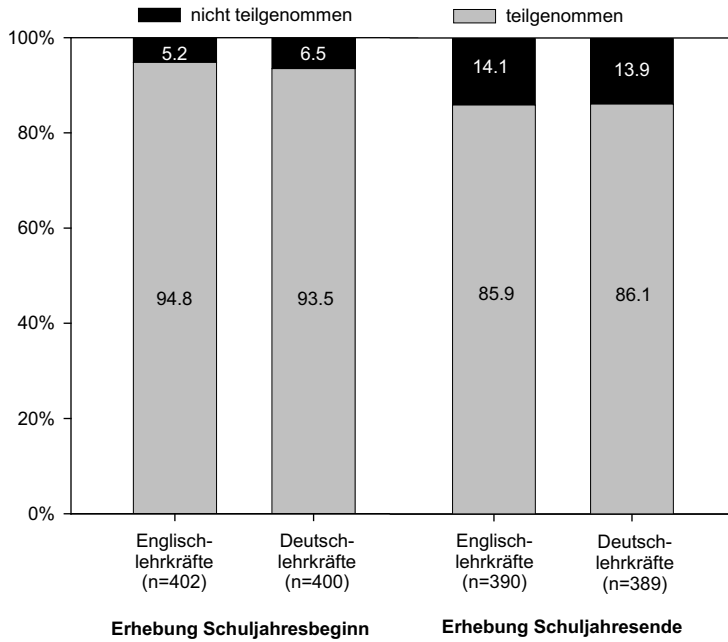


Abbildung 2.2: Teilnahmequoten für Deutsch- und Englischlehrkräfte.

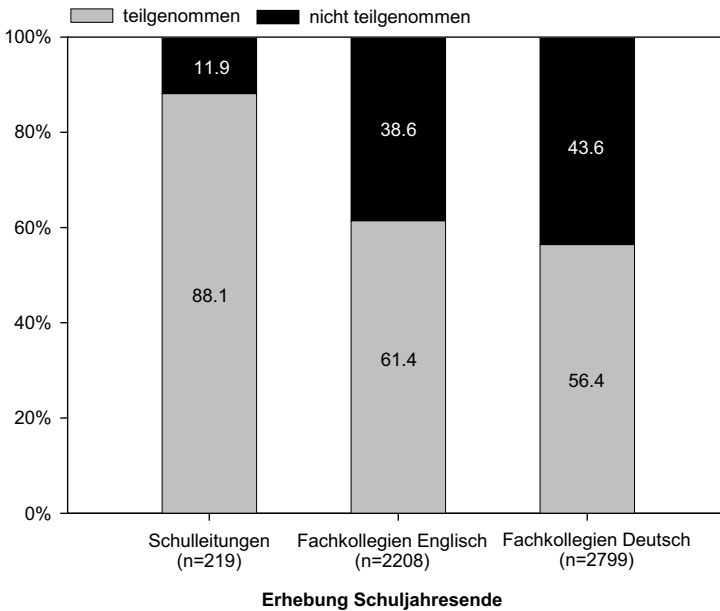


Abbildung 2.3: Teilnahmequoten für Schulleiter und Fachkollegien.

Thamar Dubberke / Birgit Harks

2.5 Zur curricularen Validität der DESI-Aufgaben: Ergebnisse eines Expertenratings

Einleitung

Die Konstruktion der DESI-Leistungstests orientierte sich eng an den Inhalten der Lehrpläne für die neunte Jahrgangsstufe der Sekundarstufe I und basierte auf detaillierten Lehrplananalysen (Beck/Klieme, 2007). Mit Hilfe von Expertenratings wurde der Frage nachgegangen, ob dieser Anspruch der curricularen Validität von unabhängigen Fachleuten gestützt wird. Dazu wurden Experten aus den Fachkommissionen der Kultusministerkonferenz (KMK) zur Erarbeitung der Bildungsstandards für die Fächer Deutsch und Englisch um eine Bewertung der DESI-Aufgaben gebeten.

Die Experten sollten beurteilen, wie wichtig es ist, die Aufgaben lösen zu können, um den Mittleren Schulabschluss zu erreichen. Diese Wichtigkeitseinschätzungen der Experten geben Aufschluss darüber, ob die in DESI gestellten Anforderungen zentrale Bestandteile der Lehrpläne sind und von den Schülerinnen und Schülern für das erfolgreiche Abschließen des Mittleren Bildungsweges erfüllt werden müssen. Darüber hinaus wurden die Experten gebeten, die DESI-Tests den von der Kultusministerkonferenz vereinbarten Bildungsstandards (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2003) zuzuordnen und für die einzelnen DESI-Aufgaben anzugeben, zu welchem Bildungsstandard ihrer Einschätzung nach die größte Entsprechung vorliegt. In den Bildungsstandards hat die Kultusministerkonferenz für verschiedene Kompetenzbereiche festgelegt, über welche Fähigkeiten Schülerinnen und Schüler bis zum Abschluss der Jahrgangsstufe zehn in der Realschule und äquivalenten Bildungsgängen verfügen sollten. Der Vergleich der zugeordneten Bildungsstandards mit den theoretischen Konzeptionen der DESI-Tests gibt Aufschluss darüber, ob die postulierten curricularen Inhalte der Aufgaben durch die Experteneinschätzungen untermauert werden.

Zum Zeitpunkt der DESI-Testentwicklung lagen die Bildungsstandards für die Fächer Deutsch und Englisch noch nicht vor und konnten dementsprechend nicht bei der testtheoretischen Konzeption berücksichtigt werden. Mit Hilfe der Expertenbeurteilung können jedoch nachträglich Aussagen darüber gemacht werden, welche Bereiche dieser von der Kultusministerkonferenz normativ gesetzten Standards durch die DESI-Aufgaben schwerpunktmäßig abgedeckt werden.

Insgesamt haben sich neun Vertreter aus dem Kreis der Fachkommissionen an dem Rating beteiligt; sechs für Aufgaben der DESI-Englischtests und drei für die Einschätzung der Aufgaben der DESI-Deutschtests. Das Expertenrating fand im Dezember 2004 und Januar 2005 statt. Aus der Gesamtheit aller Aufgaben wurden von den Autoren der DESI-Tests Aufgaben ausgewählt, die hohe Trennschärfen be-

saßen und daher die Anforderungen der Testkonzepte besonders gut widerspiegeln. Diese Auswahl an Aufgaben wurde den Experten zur Beurteilung vorgelegt.

Deutsch

Wichtigkeit der Anforderungen der DESI-Aufgaben für das Erreichen des Mittleren Schulabschlusses

Die Experten beurteilten die DESI-Deutschaufgaben als zentral für die Erreichung des Mittleren Schulabschlusses, was für eine hohe curriculare Validität der Aufgaben spricht. Abgesehen vom Rechtschreib-Test wurden die Aufgaben aller DESI-Deutschtests von den Experten im Mittel als „sehr wichtig“ eingeschätzt. Der Median des Rechtschreibtests lag bei „wichtig“. In Tabelle 2.5 (letzte Spalte) ist abgetragen, wie viel Prozent der Aufgaben der DESI-Deutschtests als „sehr wichtig“ oder „wichtig“ beurteilt wurden. Für den Rechtschreibtest lag der Anteil bei 67% und für alle anderen Deutschtests über 85%. Die Übereinstimmung der Expertenbeurteilungen für die einzelnen Subtests lag im mittleren bis hohen Bereich (Kendalls W : Spannweite = .37 bis .90, $M = .52$).

Bewertung der Aufgaben bezüglich der Bildungsstandards

Erwartungsgemäß wurden die Aufgaben der DESI-Tests Bewusstheit Deutsch, Argumentation, Rechtschreibung und Wortschatz von den Experten jeweils am häufigsten dem Kompetenzbereich „Sprache und Sprachgebrauch untersuchen“ zugeteilt. Der Test Lesekompetenz wurde hauptsächlich dem Bereich „Lesen - mit Texten und Medien umgehen“ und der DESI-Test Textproduktion am häufigsten dem Kompetenzbereich „Schreiben“ zugeordnet (vgl. Tabelle 2.5, zweite Spalte). Somit deckt DESI drei von vier Kompetenzbereichen der Bildungsstandards ab; ausgeblendet blieb lediglich der Kompetenzbereich „Sprechen und Zuhören“. Dieser Bereich beinhaltet Kompetenzen wie „zu anderen sprechen“, „vor anderen sprechen“ und „mit anderen sprechen“. Die Erfassung solcher Kompetenzen wurde bei den DESI-Deutschtests angesichts der damit verbundenen aufwendigen Testentwicklung, -durchführung und -auswertung ausgeklammert.

1 Als Maß für die durchschnittliche Übereinstimmung der Rangreihen wurde der Konkordanzkoeffizient Kendalls W berechnet. Kendalls W kann zwischen den Werten 0, bei keiner Übereinstimmung, und 1 bei vollkommener Übereinstimmung variieren (vgl. Wirtz/Casper, 2002).

Tabelle 2.5: Ergebnisse des Expertenratings für die Deutsch-Tests.

DESI-Test	KMK-Kompetenzbereich (häufigste Zuordnung)	KMK-Bildungsstandard (häufigste Zuordnung) Die Schülerinnen und Schüler...	Testkonzeption Erfasst wird:	% Aufgaben wichtig/ sehr wichtig
Sprachbewusstheit Deutsch	Sprache und Sprachgebrauch untersuchen	...können grammatische Kategorien und ihre Leistungen in situativen und funktionalen Zusammenhängen kennen und nutzen, insbesondere Tempus, Modus, Aktiv/ Passiv; Genus, Numerus, Kasus; Steigerung.	grammatische und semantische Fehler in Sätzen erkennen und Sätze in korrigierter Form aufschreiben; Konjunktivformen erkennen und korrekter Umgang mit direkter/indirekter Rede; Stilformen erkennen und schriftsprachlich umschreiben	85%
Argumentation	Sprache und Sprachgebrauch untersuchen	...können beim Sprachhandeln die Inhalts- und Beziehungsebene im Zusammenhang mit den Grundfaktoren sprachlicher Kommunikation erkennen und berücksichtigen: gelingende bzw. misslingende Kommunikation; öffentliche bzw. private Kommunikationssituationen.	Einschätzung einer Situation und erste kommunikative Reaktion; Einschätzung einer argumentativen Aussage; Produktion von Argumenten, basierend auf eigenem Wissen; argumentative Machart der Begründung	85%
Lesekompetenz	Lesen – mit Texten und Medien umgehen	...können Informationen zielgerichtet entnehmen, ordnen, vergleichen, prüfen, ergänzen.	sinntragende Wörter im Text zuverlässig finden; Inferenzen bilden und Fokus auf schwierige Stellen richten; Öffnen von Wissen und Verknüpfen v. Textstellen; innere Repräsentation wesentlicher Textaspekte haben	91%
Rechtschreibung	Sprache und Sprachgebrauch untersuchen	...kennen wichtige Regeln der Orthografie und können sie beim Sprachhandeln berücksichtigen.	Beherrschung der Schriftsprache: Lautstruktur der Wörter erkennen, entsprechend gliedern und in nachvollziehbarer Form verschriftlichen; Beherrschung der orthographischen Regeln; Beherrschung von Schreibungen, die weder durch Regeln noch Analogien erschlossen werden können	67%
Textproduktion	Schreiben	...können formalisierte lineare Texte/nichtlineare Texte verfassen: z. B. sachlicher Brief.	offizielle und/oder persönliche Briefe zu vorgegebenen Themen schreiben	91%
Wortschatz	Sprache und Sprachgebrauch untersuchen	...kennen ausgewählte Erscheinungen des Sprachwandels und können sie bewerten: z.B. Bedeutungswandel.	Umgang mit Worten aus dem Grundwortschatz, mit weniger gebräuchlichen Abstrakta und Konkreta und mit Fach- und Fremdwörtern	96%

Die Zuordnung der DESI-Deutschaufgaben zu den spezifischen Bildungsstandards stimmt im hohen Maße mit dem testkonzeptionellen Ansatz der Aufgaben überein (vgl. Tabelle 2.5). Das bedeutet, dass die postulierten curricularen Inhalte der einzelnen DESI-Tests durch die Expertenzuordnung untermauert wurden. Beispielsweise wurde dem Test zur Erfassung der Rechtschreibkompetenz, der dem testkonzeptionellen Ansatz zufolge unter anderem die Beherrschung der orthographischen Regeln erfasst, von den Experten am häufigsten der Bildungsstandard „Wichtige Regeln der Orthografie kennen und beim Sprachhandeln berücksichtigen“ zugeordnet. Beim Vergleich von konzeptionellem Ansatz eines DESI-Tests und dem am häufigsten zugeordneten Bildungsstandard muss berücksichtigt werden, dass die Bildungsstandards spezifischer formuliert sind als die testkonzeptionellen Leitideen. Dementsprechend bilden die Bildungsstandards manchmal nur einen Teil dessen ab, was mit dem jeweiligen DESI-Test laut Testkonzeption erfasst wird.

Bei der Zuordnung der Aufgaben zu den Kompetenzbereichen der Bildungsstandards waren sich alle drei Experten bei 78% der Aufgaben einig. Bei der Bewertung der Aufgaben bezüglich der Bildungsstandards wurden über die Hälfte der Aufgaben (55%) von mindestens zwei Drittel der Experten dem jeweils gleichen Standard zugeordnet. Unter Berücksichtigung der Tatsache, dass einzelne Aufgaben mehreren Standards zugeordnet werden können, hier aber nur die Zuordnung zum Bildungsstandard mit der höchsten Entsprechung analysiert wurde, ist die beschriebene Beurteilungsübereinstimmung zufriedenstellend.

Englisch

Wichtigkeit der Anforderungen der DESI-Aufgaben für das Erreichen des Mittleren Schulabschlusses

Der Großteil der Aufgaben der DESI-Englischtests wurde von den Experten als zentral für das Erreichen des Mittleren Schulabschlusses gewertet. Für sechs der sieben Englischtests lagen die mittleren Urteile der Experten bei „wichtig“ bis „sehr wichtig“, wobei für diese sechs Tests 75% bis 100% der zu beurteilenden Aufgaben als „wichtig“ oder „sehr wichtig“ bewertet wurden (vgl. Tabelle 2.6, letzte Spalte). Lediglich der Test zur Erfassung der Textrekonstruktion (C-Test) wurde von den Experten im Mittel als „eher unwichtig“ für das Erreichen des Mittleren Schulabschlusses gewertet. Beim C-Test besteht die Aufgabe der Schülerinnen und Schüler darin, einen englischen Lückentext schriftlich zu ergänzen. Der C-Test stellt ein hoch komplexes Maß der sprachlichen Gesamtkompetenz dar, der sich in der Forschung und Praxis sehr gut bewährt hat. Zur Lösung der Aufgaben ist es notwendig, Wortschatzwissen mit Grammatik-, textuellem und kontextuellem Wissen zu verschränken (vgl. Harsch/Schröder, 2007; Grotjahn, 2002). Die geringe Wichtigkeit, die den Aufgaben des C-Tests von den Experten zugesprochen wurde, lässt sich vermutlich dadurch erklären, dass das Aufgabenformat von den Experten als unterrichtsfremd angesehen und daher kritisch beurteilt wurde.

Die Übereinstimmung der Expertenratings schwankten zwischen den einzelnen Englischtests (Kendalls W : Spannweite = .17 bis .74, $M = .46$). Bei den DESI-Tests zum Lese- und Hörverstehen waren die Experten sich sehr einig (Kendalls $W = .74$ und .67), wohingegen die Wichtigkeit der Aufgaben des Tests zur Interkulturellen Kompetenz weniger einheitlich beurteilt wurde (Kendalls $W = .17$). Dies könnte eventuell mit dem Messansatz zusammenhängen, der dem DESI-Test Interkulturelle Kompetenz zugrunde liegt (vgl. Hesse/Göbel, 2007): Das Format der kritischen Interaktionssituationen (Critical Incidents) dürfte für die Leistungsprüfung im Schulalltag eher ungewöhnlich sein.

Bewertung der Aufgaben bezüglich der Bildungsstandards

Tabelle 2.6 gibt einen Überblick darüber, welchen Kompetenzbereichen und Bildungsstandards die Aufgaben der DESI-Englischtests am häufigsten zugeordnet wurden. Die Experten sahen die größte Entsprechung der DESI-Aufgaben bei drei KMK Kompetenzbereichen: Die Aufgaben zur Bewusstheit Englisch (Grammatik) und zur Textrekonstruktion (C-Test) wurden am häufigsten dem Kompetenzbereich „Verfügung über die sprachlichen Mittel“ zugeordnet. Die Fragen zur Interkulturellen Kompetenz und Bewusstheit Englisch (Pragmatik) wiesen die Experten für allem dem Kompetenzbereich „Interkulturelle Kompetenz“ zu und die Aufgaben zum Hörverstehen, Leseverstehen sowie zur Textproduktion dem Kompetenzbereich „Kommunikative Fähigkeiten“. Lediglich ein Kompetenzbereich der später als die DESI-Tests entwickelten Bildungsstandards wurde mit den durch die Experten beurteilten DESI-Aufgaben nicht abgedeckt: der Bereich „Methodenkompetenzen“ (Methoden und Arbeitstechniken zum selbstständigen Sprachenlernen)².

Die genauere Betrachtung der Beurteilung macht deutlich, dass es eine große Überschneidung zwischen den konzeptionellen Ansätzen der DESI-Tests und den zugeordneten Kompetenzbereichen der KMK gibt. Zum Beispiel wurden die DESI-Aufgaben zur Textproduktion dem Bildungsstandard „Schreiben“ zugeordnet und die Aufgaben zum Leseverstehen dem Bildungsstandard „Leseverstehen“ (in kürzeren literarischen Texten die wesentlichen Aussagen erfassen und diese zusammentragen, um eine bestimmte Aufgabe zu lösen). Dabei wurde erneut deutlich, dass die zugeordneten Bildungsstandards nur Teilaspekte der testtheoretischen Konzeptionen der DESI-Tests abbilden. Beispielsweise ordneten die Experten den C-Test am häufigsten dem Bildungsstandard „Wortschatz“ zu. Der C-Test erfasst aber – wie im vorigen Abschnitt erläutert – wesentlich komplexere Anforderungen als das Verfügen über einen umfassenden Wortschatz.

Die Analyse der Übereinstimmung der Expertenzuordnungen ergab, dass bei 63% der DESI-Aufgaben mindestens zwei Drittel der Experten in der Zuordnung

2 Einzelne Bildungsstandards, die entsprechend der Expertenbeurteilungen nicht abgedeckt wurden, waren Standards, die sich auf das Sprechen beziehen. Dies ist dadurch zu erklären, dass es in DESI einen Test zum Sprechen gibt (modifizierter SET-10 Test), dieser aber nicht in die Expertenratings aufgenommen wurde, weil er bereits auf den Gemeinsamen Europäischen Referenzrahmen für Sprachen (Europarat, 2001) abgestimmt ist (Nold/De Jong, 2007).

zu den Bildungsstandards übereinstimmen. Bezüglich der übergeordneten KMK-Kompetenzbereiche waren sich alle Experten bis auf eine Ausnahme bei allen zu beurteilenden Aufgaben einig.

Tabelle 2.6: Ergebnisse des Expertenratings für die Englisch-Tests.

DESI-Test	KMK-Kompetenzbereich (häufigste Zuordnung)	KMK-Bildungsstandard (häufigste Zuordnung) Die Schülerinnen und Schüler...	Testkonzeption Erfasst wird:	% Aufgabewichtig/ sehr wichtig
Leseverstehen	Kommunikative Fertigkeiten/ Leseverstehen	...können in kürzeren literarischen Texten (z.B. <i>Short Stories</i>) die wesentlichen Aussagen erfassen und diese zusammentragen, um eine bestimmte Aufgabe zu lösen.	Informationen (konkrete bis abstrakte) in alltäglichen Kontexten erkennen, diese verknüpfen und interpretieren können, um Hauptaussagen zu verstehen; abstrakte Informationen mit Hilfe von Inferieren impliziter Informationen verknüpfen oder inhaltlich komplexe Einzelinformationen interpretieren können	89%
Hörverstehen	Kommunikative Fertigkeiten/ Hörverstehen	...können Vorträge verstehen, wenn die Thematik vertraut und die Darstellung unkompliziert und klar strukturiert ist.	Einzelinformationen aus Kontexten alltäglicher Kommunikation (einfacher bis komplexer) hörend verstehen; konkrete Informationen beim Hören verknüpfen und ein Verständnis der Hauptaussagen entwickeln; implizite Informationen aus Gehörtem erschließen oder explizite Informationen interpretieren können	84%
Textproduktion (schriftlich)	Kommunikative Fertigkeiten/ Schreiben	...können in persönlichen Briefen Mitteilungen, einfache Informationen und Gedanken darlegen.	offizielle und/oder persönliche Briefe zu vorgegebenen Themen schreiben können	75%
Textrekonstruktion (C-Test)	Verfügung über die sprachlichen Mittel/ Wortschatz	...sind in der Lage zusätzliche lexikalische Einheiten hörend oder lesend zu verstehen oder selbstständig aus Texten zu erschließen.	allgemeine Kompetenz (genereller Sprachstand) in der Fremdsprache Englisch; Lücken in einem Text mit den korrekten Worten in der richtigen grammatikalischen Form füllen können	0%
Interkulturelle Kompetenz	Interkulturelle Kompetenz	...sind bereit, sich auf fremde Situationen einzustellen und sich in Situationen des Alltagslebens angemessen zu verhalten.	in vorgegebenen Interaktionssituationen („Critical Incidents“) relevante kulturelle Unterschiede erkennen und erfahren sowie in kulturangemessener Form denken und handeln können	91%

Tabelle 2.6: Ergebnisse des Expertenratings für die Englisch-Tests (Fortsetzung).

DESI-Test	KMK-Kompetenzbereich (häufigste Zuordnung)	KMK-Bildungsstandard (häufigste Zuordnung) Die Schülerinnen und Schüler...	Testkonzeption Erfasst wird:	% Aufgaben wichtig/ sehr wichtig
Sprachbewusstheit Grammatik	Verfügung über die sprachlichen Mittel/ Grammatik	...können Handlungen, Ereignisse und Sachverhalte als gegenwärtig, vergangen, zukünftig oder zeitlos, mehrere Geschehnisse als gleichzeitig oder aufeinanderfolgend bzw. unter Berücksichtigung von Vor- und Nachzeitigkeit erkennen und wiedergeben.	Fehler bzgl. (einfacher bis komplexer) morphologisch-syntaktischer Regelungen, syntaktischer Regelungen, Tense und Aspects erkennen und verbessern können	100%
Sprachbewusstheit Pragmatik	Interkulturelle Kompetenz	...kennen elementare spezifische Kommunikations- und Interaktionsregeln ausgewählter englischsprachiger Länder und verfügen über ein entsprechendes Sprachregister, das sie in vertrauten Situationen anwenden können.	soziokulturelle und textpragmatische Regelungen des Englischen anwenden und Verstöße oder Unregelmäßigkeiten erkennen können; sprachliches Handeln durchschauen	100%

Zusammenfassende Bewertung

Zur Überprüfung der curricularen Validität der DESI-Leistungstests wurde eine Auswahl der Aufgaben Experten der Fachkommissionen der Kultusministerkonferenz zur Erarbeitung der Bildungsstandards für die Fächer Deutsch und Englisch zur Beurteilung vorgelegt. Der Großteil der Aufgaben der DESI-Englisch- und Deutschtests wurde von den Experten als „wichtig“ oder „sehr wichtig“ für das Erreichen des Mittleren Schulabschlusses gewertet (90% der Aufgaben der Deutschtests, 82% der Aufgaben der Englischtests). Damit bestätigen die Beurteilungen der Experten die von den Testautoren intendierte Lehrplangültigkeit.

Als zusätzlicher Hinweis auf die curriculare Validität kann der Expertenvergleich der DESI-Aufgaben mit den Bildungsstandards der KMK gewertet werden. Hier zeigte sich, dass die Bildungsstandards, die die Experten den DESI-Aufgaben am häufigsten zuordneten, deutliche Überschneidungen mit den testtheoretischen Konzeptionen der DESI-Tests aufwiesen. Die postulierten curricularen Inhalte der einzelnen DESI-Tests werden also durch die Expertenzuordnung untermauert.

Ein weiterer interessanter Befund ist, dass den Expertenbeurteilungen zufolge mit den DESI-Tests zentrale durch die Kultusministerkonferenz definierte Kernkompetenzen in Deutsch und Englisch erfasst werden. Dies ist besonders bemerkenswert, da die Bildungsstandards zum Zeitpunkt der Testkonstruktion noch nicht

vorlagen und dementsprechend nicht bei der testtheoretischen Konzeption berücksichtigt wurden.

Literatur

- Beck, B./Klieme, E. (2007). Sprachliche Kompetenzen – Konzepte und Messung. Weinheim: Beltz.
- Europarat (Hrsg.) (2001): Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen. Berlin: Langenscheidt.
- Grotjahn, R. (2002). Konstruktion und Einsatz von C-Tests: Ein Leitfaden für die Praxis. In: Grotjahn, R. (Hrsg.): Der C-Test. Theoretische Grundlagen und praktische Anwendung (Bd. 4). Bochum: AKS, 211-225.
- Harsch, C./Schröder, K. (2007). Textrekonstruktion: C-Test. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen – Konzepte und Messung. Weinheim: Beltz, 212-225.
- Heller, K.A./Perleth, C. (2000): Kognitiver Fähigkeitstest für 4. bis 12. Klassen. Revision. Göttingen: Beltz.
- Hesse, H.-G./Göbel, K. (2007). Interkulturelle Kompetenz. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen – Konzepte und Messung. Weinheim: Beltz, 256-272.
- Klieme, E./Eichler, W./Helmke, A./Lehmann, R.H./Nold, G./Rolff, H.G./Schröder, K./Thomé, G./Willenberg, H. (Hrsg.) (2003): DESI: Bericht über die Entwicklung und Erprobung der Erhebungsinstrumente. Unveröffentl. Bericht. Frankfurt a.M.: DIPF.
- Nold, G./De Jong, J. H. A. L. (2007). Sprechen. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen – Konzepte und Messung. Weinheim: Beltz, 245-255.
- Schneider, W./Schlagmüller, M. (2001): Metakognitives Wissen über Textverarbeitung. Lehrstuhl für Psychologie IV, Universität Würzburg.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2003): Bildungsstandards im Fach Deutsch und in der ersten Fremdsprache für den Mittleren Schulabschluss. Beschluss der KMK vom 04.12.2003.
- Wirtz, M./Casper, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität. Göttingen: Hogrefe.

Johannes Hartig / Nina Jude / Wolfgang Wagner

3 Methodische Grundlagen der Messung und Erklärung sprachlicher Kompetenzen

Die Analyse der komplexen und umfangreichen Daten, wie sie aus Large-Scale-Assessments wie DESI resultieren, ist mit einer Reihe methodischer Herausforderungen verbunden. Verschiedene statistische Analyseverfahren, die dem aktuellen Stand der Forschungsmethodik entsprechen, kommen zum Einsatz, um eine möglichst gesicherte Interpretation der Ergebnisse zu gewährleisten. Das vorliegende Kapitel soll einen Überblick über die verschiedenen methodischen Anforderungen geben, die bei der Analyse der DESI-Daten zu berücksichtigen waren, und zugleich darüber informieren, mit welchen Verfahren gearbeitet wurde. Der erste Teil des Kapitels befasst sich mit der Messung der sprachlichen Kompetenzen, d.h. mit der Auswertung der Daten aus den Leistungstests. Hier werden kurz die verwendeten Skalierungsmodelle skizziert sowie die Schätzung der Kompetenzwerte und deren Aufbereitung für eine anschauliche Ergebnisdarstellung behandelt. Im zweiten Teil des Kapitels geht es um Fragen, die bei den weiteren Auswertungsschritten zu berücksichtigen waren, in erster Linie bei der Untersuchung von Zusammenhängen der Kompetenzen mit weiteren Schüler-, Unterrichts- und Schulvariablen. Das Kapitel kann die methodischen Grundlagen der verwendeten Analyseverfahren nicht detailliert darstellen, gibt jedoch eine Übersicht über die Verfahrenslogik der eingesetzten Techniken und verweist auf vertiefende Literatur. Methodisch vorgebildeten Leserinnen und Lesern werden Hinweise auf Verfahrensdetails gegeben, die für eine Replikation der Ergebnisse benötigt werden.

3.1 Auswertung der Leistungstests und Schätzung der Kompetenzen

Skalierungsmodelle

Die Schülerinnen und Schüler bearbeiteten in DESI eine Vielfalt verschiedener Aufgaben. Die Antworten auf diese Aufgaben erfolgten sowohl in geschlossener Form (Multiple Choice) als auch in freier schriftlicher Form. Diese Antworten wurden nach vorher definierten Kodierschemata auf ihre Korrektheit bzw. auf ihre Güte hin bewertet (vgl. Kapitel 2). Von der Güte der Testantworten der Schüler wird auf die Kompetenzen der Schüler in verschiedenen Bereichen des Deutschen und Englischen geschlossen. Je mehr Aufgaben eines Tests ein Schüler korrekt beantwortet, desto höher wird seine Kompetenz im jeweiligen Bereich eingeschätzt.

Die Kompetenztests in DESI wurden in einem Matrix-Design vorgegeben, bei dem jeder Schüler nur einen Teil der Aufgaben jedes Tests bearbeitet; zudem wurden von jedem Schüler zu Beginn und zum Ende der neunten Klasse nicht diesel-

ben, sondern unterschiedliche Aufgaben bearbeitet (vgl. Kapitel 2). Es ist daher nicht möglich, die bloße Anzahl gelöster Aufgaben in einem Bereich als einen Indikator für die jeweils interessierende Kompetenz zu verwenden, da dies die unterschiedlichen Schwierigkeiten der Aufgaben nicht berücksichtigen würde. Zur Schätzung der Schülerkompetenzen werden daher Modelle der Item-Response-Theorie (IRT) verwendet. In diesen Modellen werden die Schwierigkeiten der Testaufgaben und die Kompetenzen der getesteten Personen auf derselben Kompetenzskala beschrieben; zwischen der Kompetenz der Schüler und den erwarteten Lösungshäufigkeiten der jeweiligen Aufgaben werden statistische Zusammenhänge formuliert. Die Auswertung mit IRT-Modellen erlaubt eine gemeinsame Schätzung der Kompetenzen aller Schüler auf derselben Skala, auch wenn diese unterschiedliche Aufgaben bearbeitet haben (z.B. Rost 2004). Voraussetzung hierfür ist, dass die empirischen Daten sich mit dem verwendeten IRT-Modell hinreichend gut beschreiben lassen. Dies kann durch geeignete statistische Anpassungsmaße (*fit indices*) für einzelne Aufgaben oder ein komplettes Modell geprüft werden.

Das in DESI eingesetzte spezifische Skalierungsmodell ist ein generalisiertes Rasch-Modell (Adams/Wilson/Wang 1997; Adams/Wu 2007), welches in der Analysesoftware ConQuest implementiert ist (Wu/Adams/Wilson 1998). Innerhalb dieses Modells können sowohl dichotome (z.B. falsch/richtig) als auch ordinale Auswertungsformate (z.B. falsch/teilweise gelöst/vollständig gelöst) innerhalb desselben Tests modelliert werden. Das Raschmodell nimmt für alle Aufgaben gleich starke Zusammenhänge zwischen Kompetenz und Lösungswahrscheinlichkeit an. In Begriffen der klassischen Testtheorie bedeutet dies die Annahme gleicher Trennschärfen, in faktorenanalytischen Begriffen gleiche Faktorladungen aller Aufgaben. Diese Annahme verlangt zwar eine strengere Auswahl geeigneter Aufgaben als weniger restriktive IRT-Modelle, erlaubt aber eine konsistentere Beschreibung der Kompetenzskala unter Bezug auf die Aufgabenschwierigkeiten (Wilson 2003). Der angenommene Zusammenhang zwischen Kompetenz und Lösungswahrscheinlichkeit ist für eine zweistufige Antwortauswertung in Abbildung 3.1 graphisch veranschaulicht. Dabei ist die Schwierigkeit einer Aufgabe im Raschmodell als der Punkt auf der Kompetenzskala definiert, an dem die Aufgabe von Personen mit einer gleich hohen Kompetenz zu 50% gelöst wird. Umgekehrt können die individuellen Kompetenzen von Personen in Form von Lösungswahrscheinlichkeiten interpretiert werden: Personen, deren Kompetenz so hoch ist wie die Schwierigkeit einer Aufgabe, sollten diese Aufgabe, wenn das Modell gilt, zu 50% lösen.

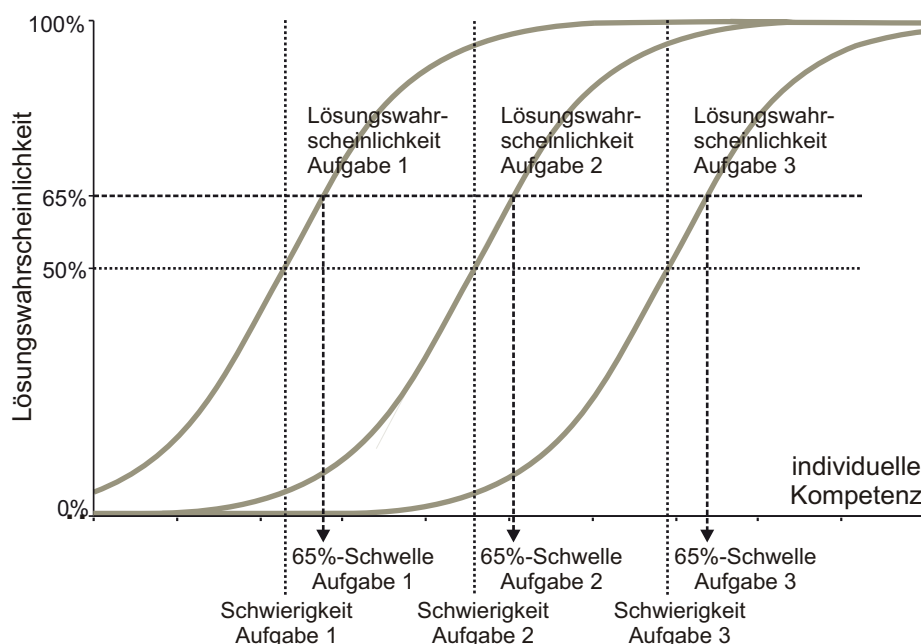


Abbildung 3.1: Veranschaulichung des im Raschmodell angenommenen Zusammenhangs zwischen individueller Kompetenz, Aufgabenschwierigkeit und Lösungswahrscheinlichkeit.

Die Beschreibung von Aufgabenschwierigkeiten und Kompetenzen auf derselben Skala und die Interpretierbarkeit in Form von Lösungswahrscheinlichkeiten stellt einen zentralen Vorteil der Auswertung mit IRT-Modellen dar und wird auch bei der Definition der Kompetenzniveaus (s.u.) herangezogen. Bei der Beschreibung der in DESI erfassten Kompetenzen interessiert, welche Schüler welche sprachlichen Anforderungen hinreichend sicher bewältigen können. Da eine Lösungshäufigkeit von 50% als Kriterium für das Beherrschen spezifischer Anforderungen zu niedrig ist, wurde in DESI eine Lösungswahrscheinlichkeit von 65% als Kriterium zur Verankerung der Aufgaben auf den Kompetenzskalen gewählt, wie sie z.B. auch in der Third International Mathematics and Science Study (TIMSS) verwendet wurde (vgl. Klieme/Baumert/Köller/Bos 2000)¹. Die im Raschmodell beschriebene Beziehung zwischen Kompetenz, Aufgabenschwierigkeit und Lösungswahrscheinlichkeit erlaubt es, diese „65%-Schwellen“ einfach zu ermitteln, indem zu der Aufgabenschwierigkeit ein konstanter Wert addiert wird. In Abbildung 3.1 ist die Bildung der 65%-Schwellen illustriert.

1 Zum Vergleich: In den PISA-Studien wird eine Lösungswahrscheinlichkeit von 62% als Kriterium für eine hinreichende Beherrschung der Aufgaben angesetzt, im US-amerikanischen National Assessment of Educational Progress (NAEP) sogar eine Lösungswahrscheinlichkeit von 80%.

Bildung von Messwerten für die erfassten Kompetenzen

Innerhalb der IRT stehen verschiedene Techniken zur Verfügung, mit der numerische Messwerte für die Kompetenzen der getesteten Personen ermittelt werden können. Eine zentrale Grundidee der verschiedenen Verfahren ist, dass die individuellen Kompetenzen so gebildet werden, dass die Wahrscheinlichkeit der tatsächlich beobachteten Antworten maximiert wird – dieses Prinzip wird als *Maximum-Likelihood-Schätzung (ML)* bezeichnet. Im Raschmodell sind die auf ML basierenden Kompetenzwerte in erster Linie eine Funktion der Anzahl gelöster Aufgaben, d.h. je mehr Aufgaben eine Person gelöst hat, desto höher ist ihre Kompetenz. Bei der Auswertung von Daten aus einem Matrixdesign spielt zusätzlich noch eine Rolle, welche Aufgabenmenge eine Person bearbeitet hatte, so dass berücksichtigt wird, ob die Menge gelöster Aufgaben auf leichteren oder schwierigeren Aufgaben basiert. Auf ML basierende Messwerte sind im Hinblick auf die Ermittlung individuell fairer Kompetenzwerte optimal, sind jedoch, wie alle individuellen Messwerte, mit einem unvermeidlichen Messfehler behaftet. In DESI wurden auf Basis der Testergebnisse als ML-basierte Kompetenzwerte *Weighted Likelihood Estimates (WLEs)*; Warm 1989) geschätzt. Diese auf der ML-Methode basierenden Schätzer gewichten die Messwerte mit der jeweils individuell verfügbaren Testinformation, um so die messfehlerbedingte Verzerrung der Populationsverteilung auszugleichen. Die individuellen Messwerte sind jedoch ebenso messfehlerbehaftet wie ungewichtete ML-Schätzer.

Das Ziel von Large-Scale-Assessments wie DESI ist jedoch nicht die Messung von Kompetenzen der einzelnen getesteten Schüler, wie dies z.B. bei Abschlussprüfungen oder Vergleichsarbeiten der Fall ist. Es geht vielmehr darum, die Verteilungen der Kompetenzen in der interessierenden Grundgesamtheit (d.h. aller Neuntklässler) und in interessierenden Untergruppen (z.B. differenziert nach Geschlecht oder Bildungsgang) zu beschreiben. Diese beiden Zielsetzungen – Individualdiagnostik und Populationsbeschreibung – stehen insofern in einem gewissen Konflikt, als individuelle Messwerte für die erfassten Kompetenzen immer mit einem Messfehler behaftet sind. Dies führt dazu, dass ML-basierte Messwerte zwar für die individuelle Kompetenzbeschreibung optimal sind, nicht jedoch, um die Verteilung der Kompetenzen und Zusammenhänge mit weiteren Personenmerkmalen in der interessierenden Grundgesamtheit unverzerrt zu beschreiben (Mislevy/Beaton/Kaplan/Sheehan 1992). Um dem Ziel von DESI als repräsentativer Studie gerecht zu werden und die Verteilung sprachlicher Kompetenzen in der Gesamtheit der Neuntklässler sowie die Zusammenhänge mit Schüler- und Unterrichtsmerkmalen optimal zu beschreiben, wird in DESI, wie auch in anderen Large-Scale-Assessments, von der Plausible-Value-Technik Gebrauch gemacht (Mislevy/Beaton/Kaplan/Sheehan 1992; Adams/Wu/Carstensen 2007). Hierbei werden bei der Schätzung der Kompetenzen nicht nur die Antworten auf die Testaufgaben herangezogen, sondern auch Eigenschaften der untersuchten Personen wie z.B. Geschlecht, Erstsprache oder Bildungsgang. Noch vor der Bestimmung individueller Kompetenzwerte werden die Zusammenhänge

zwischen diesen *Hintergrundvariablen* und der zu messenden Kompetenz ermittelt, z.B. ob Mädchen durchschnittlich höhere Werte haben als Jungen. Diese Zusammenhänge, die nicht auf der Schätzung individueller Werte basieren, werden im so genannten *Hintergrundmodell* auf latenter Ebene und damit messfehlerfrei geschätzt. Als Hintergrundvariablen wurden in DESI alle Variablen aus allen eingesetzten Fragebögen (Schüler, Eltern, Lehrer, Schulleiter und Fachkollegien; vgl. Kapitel 2) sowie die individuellen Leistungen im Kognitiven Fähigkeitstest (KFT, Heller/Perleth 2000) einbezogen. Zusätzlich wurde, um der für Schulleistungsdaten charakteristischen Stichprobenstruktur gerecht zu werden (s.u.), der mittlere Anteil gelöster Aufgaben in einer Klasse im jeweils interessierenden Test in das Hintergrundmodell einbezogen. Um die große Menge dieser sowohl auf individueller Ebene als auch auf Klassen- und Schulebene erfassten Daten in das Hintergrundmodell einbeziehen zu können, wurden die Daten mittels Hauptkomponentenanalysen zu unkorrelierten Faktorwerten zusammengefasst, die insgesamt 99.9% der Variation in den ursprünglichen Variablen erfassen. Das so konstruierte Hintergrundmodell berücksichtigt praktisch alle linearen Zusammenhänge, die zwischen den in DESI erfassten Kompetenzen und anderen Variablen bestehen. Zusätzlich wurden für die Faktoren Bildungsgang, Geschlecht und Erstsprache auch die Wechselwirkungen erster und zweiter Ordnung aufgenommen.

Anschließend wurden für die untersuchten Schülerinnen und Schüler *Plausible Values (PVs)* erzeugt; zusätzlich wurden auch die jeweiligen WLEs für die erfassten Kompetenzen ermittelt. Die PVs berücksichtigen sowohl die Menge gelöster Aufgaben als auch die Eigenschaften der untersuchten Personen. Wenn im ersten Schritt z.B. ein Kompetenzvorsprung für Mädchen ermittelt wurde, wird dieser bei der Schätzung der PVs dahingehend berücksichtigt, dass Mädchen etwas höhere Werte zugewiesen bekommen als Jungen mit der gleichen Anzahl gelöster Aufgaben. Dieses Vorgehen ist offensichtlich nicht zur Schätzung fairer individueller Messwerte konzipiert, erlaubt jedoch eine unverzerrte Schätzung der Kompetenzverteilungen in der zugrunde liegenden Population sowie der Zusammenhänge zwischen individuellen Hintergrundmerkmalen und der untersuchten Kompetenz. Es ist ein Charakteristikum der Plausible-Value-Technik, die auf der multiplen Imputation fehlender Werte basiert (Schafer 1997), dass für jede interessierende Kompetenz mehrere – im Fall von DESI fünf – PVs generiert werden (vgl. auch Walter 2006). Analysen auf Basis der PVs werden mit jedem PV wiederholt, und die Streuung zwischen den fünf Analyseergebnissen wird bei der Einschätzung der statistischen Bedeutsamkeit der Ergebnisse berücksichtigt.

Im DESI-Datensatz lagen also für jede sprachliche Kompetenz zwei Typen von Kompetenzschätzungen vor, nämlich WLEs und PVs. Je nach Art der Analyse kommen beide Werte zur Anwendung. Die weitaus meisten Analysen, bei denen die Zusammenhänge zwischen den Schülerkompetenzen und weiteren individuellen, Unterrichts- oder Schulvariablen von Interesse sind, werden mit den PVs durchgeführt, da diese die besten Schätzungen für die Populationsverteilungen der Kompetenzen liefern. Eine simultane Erzeugung von PVs für alle Kompetenzen in einer gemeinsa-

men mehrdimensionalen Analyse war in DESI durch die große Anzahl verschiedener Kompetenzen technisch nicht möglich. Für die Untersuchung von Zusammenhängen zwischen den verschiedenen sprachlichen Kompetenzen werden daher die jeweiligen WLEs verwendet, da die PVs aus separaten Skalierungen stammen und damit die Zusammenhänge zwischen den verschiedenen Kompetenzen nicht korrekt wiedergeben. Die mittels WLEs geschätzten Zusammenhänge sind zwar messfehlerbehaftet, dies kann jedoch durch die Modellierung mit latenten Variablen berücksichtigt werden (vgl. Kapitel 18; Hartig/Jude 2005). Auch die Rückmeldungen an die Lehrkräfte der in DESI untersuchten Schulklassen (vgl. auch Kapitel 4) wurden auf Basis der WLEs vorgenommen.

Für alle in DESI eingesetzten Tests wurden jeweils separat individuelle Werte in Form von WLEs und PVs gebildet. Zusätzlich wurde auf Basis der Tests, die zum Ende der neunten Jahrgangsstufe eingesetzt worden waren, jeweils ein Gesamtwert für Deutsch und Englisch gebildet. Hierzu wurden die Daten der einzelnen Tests in einer gemeinsamen Skalierung – d.h. wie Antworten auf einen einzigen großen Test – analysiert. Die einzige Ausnahme stellte hierbei der C-Test zur Textrekonstruktion dar. Diese Lückentexte waren zur Bildung der Kompetenzwerte auf Ebene der einzelnen Lücken skaliert worden (Harsch/Schröder 2007). Da der Textrekonstruktions-Test hierdurch sehr viele „Aufgaben“ enthält, wäre er bei der Bildung des Gesamtwertes gegenüber den übrigen, kürzeren Tests übergewichtet worden. Um dies zu vermeiden und dennoch die wesentlichen Informationen über Leistungsunterschiede in Textrekonstruktion im Englisch-Gesamtwert zu berücksichtigen, wurde die Leistung aus den einzelnen Lückentexten zu Kategorien (0 bis 5 richtig gefüllte Lücken, 6 bis 10 richtig gefüllte Lücken, usw.) zusammengefasst und die einzelnen Texte als mehrstufige Items behandelt. Sowohl für Deutsch als auch für Englisch wurden als Schätzung für die Gesamtkompetenz der Schüler in der jeweiligen Sprache WLEs und PVs geschätzt. Diese Gesamtwerte für Deutsch und Englisch werden dort als abhängige Variablen verwendet, wo eine Differenzierung nach den einzelnen Kompetenzen aus inhaltlicher Sicht nicht angezeigt ist und mit einer Verwendung der Gesamtwerte eine übersichtlichere Ergebnislage erreicht werden kann. Beispiele sind die Effekte von Unterrichts- und Schuleigenschaften (vgl. Kapitel 26 bis 34).

Zu Beginn der neunten Jahrgangsstufe wurde lediglich eine Auswahl der DESI-Tests eingesetzt. Es ist daher nicht möglich, für diesen Zeitpunkt einen Gesamtwert zu bilden, der mit der Gesamtleistung am Ende der neunten Jahrgangsstufe vergleichbar ist; aus diesem Grund wurden auch keine PVs für eine Gesamtleistung zu Beginn der neunten Jahrgangsstufe gebildet. Für einige Analysen ist es jedoch erstrebenswert, das Leistungsniveau der Schülerinnen und Schüler zu Beginn des Untersuchungszeitraums als Kontrollvariable zu verwenden. Um dies zu ermöglichen, wurde als Indikator für das Ausgangsniveau der sprachlichen Kompetenzen zu Beginn der neunten Jahrgangsstufe für Deutsch und Englisch jeweils der Mittelwert der WLEs für die zu Beginn der Untersuchung eingesetzten Tests gebildet.

Normierung der Kompetenzskalen

Die Schätzungen für individuelle Kompetenzwerte werden in Analysen mit dem in DESI verwendeten IRT-Modell auf einer so genannten Logit-Skala dargestellt. Diese Werte liegen typischerweise in einem numerisch relativ kleinen Wertebereich um null (z.B. -3 bis 3), die genaue Streuung hängt von den Messeigenschaften des jeweiligen Tests ab. Um anschaulichere Werte für die in DESI gemessenen Kompetenzen zu erhalten, wurden die für die weiteren Analysen und die deskriptiven Ergebnisdarstellungen verwendeten Werte so normiert, dass der Mittelwert der Grundgesamtheit *am Ende der neunten Jahrgangsstufe* auf 500 Punkte und die Standardabweichung auf 100 Punkte gesetzt wurde. Ein Schüler mit einem Wert von 550 Punkten für eine bestimmte Kompetenz würde in diesem Bereich also eine halbe Standardabweichung über dem Mittelwert aller Neuntklässler liegen. Die Skalierung mit einem Mittelwert von 500 und einer Standardabweichung von 100 Punkten ist an die Normierung der Kompetenzwerte in den PISA-Studien angelehnt (z.B. OECD 2001, 2004). Es ist jedoch zu beachten, dass in PISA der internationale Mittelwert der teilnehmenden OECD-Länder als Referenz verwendet wurde, und der Mittelwert der deutschen Schüler z.B. in Lesen in PISA 2000 ca. 480 Punkte betrug. In DESI, wo kein internationaler Vergleich vorgenommen wird, entsprechen 500 Punkte dem Mittelwert der deutschen Neuntklässler. Die numerischen Werte auf Basis dieser Normierung sind nicht mit den Werten aus den PISA-Studien vergleichbar. Die Verteilung am Ende der neunten Jahrgangsstufe wurde als Referenz für die Normierung gewählt, da zum Ende der neunten Jahrgangsstufe fast alle Tests eingesetzt wurden, zu Beginn der neunten Jahrgangsstufe nur eine Auswahl (vgl. auch Kapitel 2). Die Normierung der Messwerte zu Beginn der neunten Jahrgangsstufe wurde ebenfalls auf Basis der Verteilungen zum Ende der neunten Jahrgangsstufe vorgenommen. Dies führt dazu, dass sich für die zu Beginn der neunten Jahrgangsstufe gemessenen Kompetenzen normierte Mittelwerte kleiner als 500 ergeben, wenn innerhalb der neunten Jahrgangsstufe ein Zuwachs stattgefunden hat. Eine Ausnahme von diesem Vorgehen bildet der Test für Englisch Sprachbewusstheit im Bereich Soziopragmatik (Nold/Rossa 2007), der nur zu Beginn der neunten Jahrgangsstufe eingesetzt worden war. Die Messwerte für diesen Test wurden auf Basis der Werteverteilung zu Beginn der neunten Jahrgangsstufe normiert.

Die Normierung der Kompetenzskalen erfolgte auch für die Itemschwierigkeiten. Die 65%-Schwellen (s.o.) für alle Items wurden in der gleichen Weise transformiert wie die Messwerte für die jeweiligen Tests. Aufgabenschwierigkeiten und Kompetenzen werden so weiterhin auf derselben Skala beschrieben. Die Transformation der 65%-Schwellen führt dazu, dass zum Beispiel eine Aufgabe mit einer transformierten Schwierigkeit von 500 Punkten von Schülerinnen und Schülern mit einer durchschnittlichen Kompetenz zu ca. 65% gelöst wird. Eine Aufgabe mit einer transformierten Schwierigkeit von 600 Punkten wird nur von Schülerinnen und Schülern, deren Kompetenz eine Standardabweichung über dem Mittelwert liegt, mit einer Häufigkeit von 65% gelöst. In den Kapiteln zu den Deutsch- und

Englischkompetenzen (Kapitel 5-15) werden, soweit auf Aufgabenschwierigkeiten Bezug genommen wird, immer diese transformierten Aufgabenschwierigkeiten verwendet.

Schätzung von Kompetenzzuwächsen

In DESI wurden Schülerkompetenzen in ausgewählten Bereichen zu Beginn und am Ende der neunten Jahrgangsstufe erfasst. Ein zentrales Anliegen der Studie ist es, die Zuwächse, die in den verschiedenen sprachlichen Kompetenzen zwischen diesen Zeitpunkten erzielt werden, einzuschätzen sowie mögliche Prädiktoren für Unterschiede in den Zuwächsen zu untersuchen. In der IRT gibt es verschiedene Techniken, mit Daten aus mehreren Messzeitpunkten umzugehen (Rost 2004; Meiser/Stern/Langeheine 1998). Wenn Schätzungen der Kompetenzen zu mehreren Zeitpunkten miteinander verglichen werden sollen, ist es wichtig, dass die Schwierigkeiten der Aufgaben zu den verschiedenen Zeitpunkten im Modell gleich gesetzt werden. Eine Technik, dies bei der Schätzung individueller Kompetenzwerte für mehrere Messzeitpunkte zu erreichen, ist die Bildung sogenannter *virtueller Personen*. Die Messungen zum zweiten und weiteren Zeitpunkten werden hierbei so behandelt, als ob zusätzliche Personen den Test bearbeitet hätten (vgl. Rost 2004). Während diese Methode für ML-basierte Kompetenzschätzungen geeignet ist, führt sie bei der Schätzung von PVs zu stark verzerrten Verteilungsschätzungen, da die Abhängigkeiten zwischen den Antworten derselben Personen nicht berücksichtigt werden (Hartig/Kühnbach 2006). Diesen Abhängigkeiten wird Rechnung getragen, wenn die Daten aus den verschiedenen Messzeitpunkten in einem mehrdimensionalen Modell analysiert werden, in dem jeder Zeitpunkt als eine separate Kompetenzdimension behandelt wird. Durch ein Gleichsetzen der Aufgabenschwierigkeiten zwischen den Zeitpunkten kann erreicht werden, dass die Kompetenzschätzungen zwischen den Zeitpunkten vergleichbar sind. Werden auf Basis eines solchen mehrdimensionalen Modells PVs generiert, ermöglichen diese im Unterschied zu ML-basierten Kompetenzwerten eine messfehlerbereinigte Schätzung der Zusammenhänge sowohl zwischen den beiden Zeitpunkten als auch zwischen anderen Variablen und dem Kompetenzzuwachs zwischen den Zeitpunkten (Hartig/Kühnbach 2006; Hartig/Frey 2006).

In DESI wurden für die Tests, die zu beiden Messzeitpunkten in der neunten Jahrgangsstufe eingesetzt wurden, zunächst in einer Skalierung mit virtuellen Personen, d.h. in einem eindimensionalen Modell unter Einbezug aller vorhandenen Daten, die Aufgabenschwierigkeiten geschätzt. In einem zweiten Schritt wurde unter Verwendung dieser Aufgabenschwierigkeiten ein zweidimensionales Modell geschätzt, in dem die Kompetenz zu Beginn der neunten Jahrgangsstufe als eine, die Kompetenz am Ende der neunten Jahrgangsstufe als zweite latente Dimension behandelt wurde. Das Hintergrundmodell wurde wie oben beschrieben konstruiert. Die auf Basis dieses zweidimensionalen Modells generierten PVs wurden als Kompetenzschätzungen für die beiden Messzeitpunkte in der neunten

Jahrgangsstufe verwendet. Zusätzlich wurden die Differenzen zwischen den jeweils korrespondierenden PVs der beiden Zeitpunkte gebildet. Diese PV-Differenzen dienen als Schätzungen für den Kompetenzzuwachs zwischen Anfang und Ende des Schuljahres.

Definition der Kompetenzniveaus

Die für DESI vorgenommenen Kompetenzschätzungen sind quantitative Werte auf kontinuierlichen Skalen, die eine sehr feine Abstufung von Kompetenzunterschieden erlauben. Bei der Untersuchung von Zusammenhängen der Kompetenzen mit anderen Variablen ist diese Differenzierung von großem Vorteil. Es ist hingegen schwer, die Bedeutung unterschiedlich hoher Kompetenzwerte bezogen auf die konkreten inhaltlichen Anforderungen der Testaufgaben in dieser Differenziertheit anschaulich zu beschreiben (Beaton/Allen 1992). Für die Interpretation der DESI-Ergebnisse ist es jedoch sehr wichtig, welche konkreten Anforderungen im sprachlichen Bereich die untersuchten Schüler bewältigen können – es soll eine kriterienorientierte Interpretation der beobachteten Testleistung ermöglicht werden (z.B. Goldhammer/Hartig 2007; Rauch/Hartig 2007). Um eine solche anschauliche Interpretation der gemessenen Kompetenzen zu erreichen, werden die kontinuierlichen Kompetenzskalen in Abschnitte unterteilt (z.B. Beaton/Allen 1992, Klieme/Baumert/Köller/Bos 2000, OECD 2001, 2004). Diese Abschnitte werden als Kompetenzniveaus oder Kompetenzstufen² bezeichnet. Für jedes Kompetenzniveau wird dann eine anschauliche Beschreibung der Anforderungen vorgenommen, welche die Schüler, deren Leistungen im entsprechenden Abschnitt der Skala liegen, bewältigen können. Entscheidend bei der Definition der Kompetenzniveaus ist, an welchen Stellen der Skala die Schwellen zwischen den Niveaus gesetzt werden.

In DESI wurden zur Setzung der Schwellen zwischen den Kompetenzniveaus systematisch definierte Merkmale der Testaufgaben verwendet. Diese Aufgabenmerkmale, die für jeden der einzelnen Tests im vornherein definiert wurden, beschreiben spezifische Anforderungen, hinsichtlich derer sich leichtere von schwereren Aufgaben unterscheiden sollten. Die theoretisch angenommenen Zusammenhänge der Aufgabenmerkmale mit den tatsächlich ermittelten Aufgabenschwierigkeiten können in Regressionsanalysen untersucht werden. Aus diesen Analysen lassen sich auch erwartete Schwierigkeiten für bestimmte Kombinationen von Aufgabenmerkmalen ableiten. Diese erwarteten Schwierigkeiten wurden in DESI herangezogen, um die Schwellen zwischen den Kompetenzniveaus zu definieren. Hierdurch ist es möglich, die gemessenen Kompetenzen bezogen auf die sprachlichen Anforderungen

2 Angesichts möglicherweise irreführender Konnotationen des Stufenbegriffs (vgl. Helmke/Hosenfeld 2004) wurde für DESI der Begriff des Kompetenzniveaus dem der Kompetenzstufe vorgezogen. Die Kompetenzniveaus in DESI bezeichnen jedoch dasselbe wie Kompetenzstufen im Kontext anderer Studien, nämlich Abschnitte auf kontinuierlichen Kompetenzskalen, die mit dem Ziel einer kriteriumsorientierten Beschreibung der erfassten Kompetenzen gebildet werden.

in den verschiedenen Kompetenzbereichen zu beschreiben. Das generelle Vorgehen zur Definition der Kompetenzniveaus in DESI ist ausführlich bei Hartig (2007) beschrieben, die Aufgabenmerkmale für die einzelnen Kompetenzen in den jeweiligen Kapiteln in Beck/Klieme (2007).

3.2 Analyse von statistischen Zusammenhängen

Umgang mit fehlenden Werten

Es ist für Large-Scale-Assessments wie DESI typisch, dass die erfassten Daten fehlende Werte aufweisen. Zum Teil werden Daten aus ökonomischen Gründen von vornherein nur unvollständig erhoben (*missing by design*), wie z.B. die im Matrix-Design vorgegebenen Testaufgaben oder die fachspezifischen Fragen in den Schülerfragebögen. Weitere Daten fehlen aufgrund unvollständiger Beteiligungen an den Schüler- und Elternbefragungen (zu Erhebungsdesign und Beteiligungsdaten s. Kapitel 2). Fehlende Werte aufgrund unterschiedlicher Beteiligungsdaten fehlen in der Regel nicht zufällig, sondern hängen systematisch mit bestimmten Merkmalen von Schülern und Eltern zusammen; so sind z.B. die Beteiligungsdaten an der Elternbefragung bei Eltern mit nicht deutscher Erstsprache geringer als bei deutschsprachigen Eltern³. Bei Analysen mit derartigen systematisch fehlenden Werten würde ein Ausschluss der Fälle mit fehlenden Werten aus der Analyse zu einer Verzerrung der Analyseergebnisse führen (z.B. Wolf 2006). Um unverzerrte Ergebnisse zu erhalten, können die fehlenden Werte mit einer *Multiplen Imputation* geschätzt werden (Schafer 1997). Hierbei werden vorhandene Informationen über Zusammenhänge zwischen den interessierenden Variablen sowie vorhandene Informationen über die Fälle mit teilweise fehlenden Werten zur Schätzung dieser fehlenden Werte herangezogen.

Für die Sprachkompetenzen erfolgte eine solche Schätzung der fehlenden Werte bei der oben beschriebenen Erzeugung der PVs, da die Plausible-Values-Technik auf der Multiplen Imputation basiert. Jeder Schüler und jede Schülerin, für den / die zu einem der beiden Messzeitpunkte in der neunten Jahrgangsstufe ein bearbeiteter Test vorlag, wurde als Teilnehmer/-in an der Studie definiert. Für diese $N = 10543$ Fälle wurden auf Basis des oben beschriebenen Hintergrundmodells PVs für die erfassten Kompetenzen generiert, auch wenn der jeweils interessierende Test nicht bearbeitet wurde. Auf diese Weise wurde vor allem die Datenbasis für die Schätzung der Kompetenzzuwächse erweitert, da PVs für beide Zeitpunkte auch für Fälle generiert wurden, die nur zu einem der beiden Zeitpunkte an der Testung teilgenommen hatten. Der Anteil der durch die Bildung der PVs geschätzten fehlenden Werte beträgt für alle Tests unter 10%. Dieser Anteil ist vergleichsweise gering, selbst bei ca. 40% fehlenden Werten in einzelnen Variablen kann eine Multiple Imputation noch zu bes-

3 Eine Regression mit den Variablen Erstsprache (Deutsch vs. nicht deutsch), Migrationsstatus, höchster Bildungsabschluss der Eltern und der Deutschleistung (inklusive aller Interaktionsterme) erklärt etwa 8% der Varianz im Fehlen des Elternfragebogens.

seren Analyseergebnissen führen als ein Ausschluss der Fälle aus der Analyse (z.B. Schafer/Olsen 1998).

Auch für eine zentrale Hintergrundvariable in DESI, den höchsten sozioökonomischen Status in der Familie (*Highest International Socio-Economic Index*, HISEI) nach Ganzeboom u.a. (1992), wurde eine Schätzung fehlender Werte vorgenommen. Dieser Index basiert auf Elternangaben. Für etwa 62% der Schülerinnen und Schüler lagen die dafür erforderlichen Angaben vor. Für weitere 29%, für die ein beantworteter Schülerfragebogen vorlag, wurde der Index geschätzt. Die Imputation erfolgte mit dem Programm *IVEWare* (Raghunathan/Solenberger/Van Hoewyk 2002). Auch hier wurden – analog zu den Plausible Values – fünf Werte generiert. Dazu wurde neben verschiedenen Schülerangaben, die in der gesamten Stichprobe erhoben wurden, insbesondere auf Angaben zum Beruf bzw. Schulabschluss der Eltern zurückgegriffen. Zur Berücksichtigung der Mehrebenenstruktur der Daten wurden die auf Klassenebene aggregierten Variablen (Klassenmittelwerte) als Faktorwerte (Hauptkomponentenanalyse, 24 Faktoren mit etwa 90% Varianzaufklärung) in die Imputation einbezogen.

Eine polytome Regression – inklusive einiger wichtiger Interaktionsterme – auf der Basis der Schülerangaben zum Beruf und dem Schulabschluss der Eltern erwies sich als optimales Verfahren bezüglich der Varianzaufklärung des HISEI. Nachteil dieses Verfahrens ist der verglichen mit der linearen Regression extrem hohe Rechenaufwand. Deshalb wurden fehlende Werte in den übrigen Variablen mittels linearer Regression imputiert. Erst in einem zweiten Schritt wurde dann der HISEI auf der Basis eines Datensatzes, bei dem – mit Ausnahme der Angaben zum Beruf bzw. Schulabschluss – alle Variablen bereits vollständig imputiert waren, mit Hilfe einer polytomen Regression imputiert.

Das gesamte Verfahren wurde anhand der vorliegenden Daten zum HISEI überprüft, indem analog zu den fehlenden Elternangaben Daten „gelöscht“ und anschließend imputiert wurden. Es zeigte sich, dass die Ergebnisse von Zwei-Ebenen-Analysen auf Basis der „echten“ bzw. der imputierten Daten sowohl bezüglich der Varianzen als auch hinsichtlich der Signifikanztests auf beiden Ebenen weitgehend vergleichbar sind. Zusätzlich wurden ausgewählte Zusammenhänge auf der Basis der gesamten Daten (mit fehlenden vs. mit imputierten HISEI-Werten) überprüft⁴. Auch hier zeigten sich vergleichbare Ergebnisse.

Bei allen übrigen Variablen und Skalen wurde auf eine solche aufwändige Imputation verzichtet. Einige Skalen aus den Schülerfragebögen wurden zwar nur bei der Hälfte der Schülerinnen und Schüler der jeweiligen Klassen erhoben, so dass auch dort eine Imputation auf den ersten Blick nahe gelegen hätte, dies betrifft allerdings nur die Zahl der fehlenden Werte. Da es sich hier um designbedingte „rein zufällig fehlende Werte“ (*missing completely at random*, MCAR) handelt, sind die Schätzer für Mittelwerte, Standardabweichungen und Zusammenhänge ohnehin un-

4 Dazu wurden mit Hilfe des Programms Mplus 3.12 Korrelationsmatrizen auf zwei Ebenen mit der so genannten *missing at random maximum likelihood missing data*-Methode (MAR ML) geschätzt.

verzerrt (z.B. Wolf 2006). Bei Zusammenhangsanalysen wurde stattdessen auf das in der Analysesoftware Mplus implementierte Verfahren für Analysen mit fehlenden Werten zurückgegriffen.

Berücksichtigung der Stichprobenstruktur bei Signifikanztests und Standardfehlern

Ein wesentliches Ziel von DESI ist es, die Sprachkompetenzen in verschiedenen Gruppen von Schülern zu beschreiben, z.B. in verschiedenen Bildungsgängen oder mit verschiedenen sprachlichen Hintergründen. Wenn Unterschiede zwischen Gruppen zu beobachten sind, stellt sich die Frage, ob diese Unterschiede in der untersuchten Stichprobe nur zufällig zustande gekommen sind, oder ob sie auf die Gesamtheit der Neuntklässler übertragbar sind. Diese Frage wird mit statistischen Signifikanztests beantwortet. Hierbei wird geprüft, mit welcher Wahrscheinlichkeit ein vorgefundenes Ergebnis aus einer Population stammt, in welcher der interessierende Zusammenhang *nicht* vorliegt, d.h. die so genannte Nullhypothese gilt. Wenn diese Wahrscheinlichkeit unter einem vorher definierten Signifikanzniveau (üblicherweise 5%) liegt, wird die Nullhypothese verworfen und es wird davon ausgegangen, dass der Zusammenhang auch in der Grundgesamtheit, aus der die Stichprobe gezogen wurde, von null verschieden ist. Ein in diesem Zusammenhang wichtiges Konzept ist der *Standardfehler* statistischer Kennwerte. Dieser gibt die Genauigkeit an, mit der ein Wert – z.B. ein Mittelwert oder ein Korrelationskoeffizient – auf Basis der Stichprobendaten geschätzt werden kann. Je kleiner der Standardfehler, desto zuverlässiger kann von den Stichprobendaten auf die Population geschlossen werden. Mit Hilfe des Standardfehlers kann ein *Konfidenzintervall* definiert werden, das einen Bereich von ungefähr \pm zwei Standardfehlern unter- und oberhalb des Mittelwertes umfasst und beschreibt, in welchem Wertebereich wahrscheinlich der interessierende Wert in der untersuchten Grundgesamtheit liegt. Je größer der Standardfehler und je breiter dadurch das Konfidenzintervall, desto ungenauer ist die Schätzung des jeweils interessierenden Kennwertes.

Bei der Schätzung von Standardfehlern und Signifikanztests werden in den Sozialwissenschaften in der Regel bestimmte Grundannahmen über das Verhältnis von Stichprobe und Grundgesamtheit gemacht. Die meisten häufig verwendeten Signifikanztests setzen voraus, dass die Fälle in der Stichprobe zufällig aus der Grundgesamtheit gezogen wurden. Diese Voraussetzung ist in DESI und vielen anderen Schulleistungsstudien nicht gegeben. Die interessierende Grundgesamtheit sind alle Schülerinnen und Schüler innerhalb Deutschlands in der neunten Klasse, gezogen wurden jedoch zunächst Schulen und dann Schulklassen innerhalb der Schulen (vgl. Kapitel 2). Innerhalb der Schulklassen wurden dann jeweils alle Schüler untersucht. Die resultierende Stichprobe hat eine so genannte *Klumpen-* oder *Clusterstruktur*, da die Fälle in Gruppen vorliegen, die sich untereinander ähnlicher sind (nämlich Schüler derselben Klassen) als zwischen den Gruppen (d.h. Schüler verschiedener Klassen). Wenn Daten mit einer solchen Clusterstruktur mit statisti-

schen Standardverfahren analysiert werden, führt dies zu einer Unterschätzung der Standardfehler, und Zusammenhänge werden zu schnell als signifikant betrachtet.

Es gibt zwei gängige Methoden, mit der Clusterstruktur von Stichproben angemessen umzugehen. Zum einen können die Standardfehler und Signifikanzen durch so genannte *Resampling*-Verfahren unter Berücksichtigung der Stichprobenstruktur ermittelt werden, wobei auch der Stichprobenplan und die Gewichtung einbezogen werden. Die statistischen Verfahren sind hierbei einfache Standardanalysen (z.B. Deskriptive Statistiken, Mittelwertsvergleiche, Korrelationsanalysen). Die entsprechenden Analysen werden jedoch mit systematisch variierender Gewichtung wiederholt, und die Variation der Ergebnisse zwischen den wiederholten Analysen wird in die Standardfehler und Signifikanztests mit einbezogen (Brick/Morganstein/Valliant 2000). Für DESI wurde zur Schätzung der Standardfehler und Signifikanzen bei einfachen Zusammenhangsanalysen die Software WesVar (Westat 2000) verwendet, die dieses leistet. Als Resampling-Methode wurde die *Balanced Repeated Replication* (BRR) nach Fay (1989) mit einem Perturbationsfaktor von 70% ($K = 0.3$; vgl. Westat 2000) verwendet. Diese Analysesoftware erlaubt neben der Berücksichtigung der Stichprobenstruktur auch simultan die Berücksichtigung der Streuungen zwischen den Plausible Values (s.o.).

Eine zweite Möglichkeit der Berücksichtigung der Clusterstruktur ist die Anwendung von *Mehrebenenmodellen* (z.B. Hox 2003; auch *Hierarchisch Lineare Modelle*). Hierbei wird der Umstand, dass die untersuchten Schüler in kompletten Klassen rekrutiert wurden, dadurch berücksichtigt, dass die Schulklasse als eine separate Analyseebene in die statistischen Modelle mit einbezogen wird. Die Unterschiede in den gemessenen Kompetenzen werden hierbei in Varianz *zwischen Klassen* („Ebene 2“) und in Varianz *zwischen Schülern* innerhalb von Klassen („Ebene 1“) aufgeteilt. Zusätzlich kann berücksichtigt werden, dass sich statistische Zusammenhänge zwischen verschiedenen Klassen unterscheiden können. Der Einsatz von Mehrebenenmodellen ist insbesondere dann angezeigt, wenn Variablen auf Klassenebene, wie z.B. Unterrichtseigenschaften, im Mittelpunkt des Interesses stehen. In DESI kamen Mehrebenenmodelle vor allem bei Analysen zu Unterrichtsmerkmalen und Schuleigenschaften zum Einsatz. Bei den Analysen zu Schuleigenschaften stellt die Schule die zweite Ebene dar, bei allen anderen Mehrebenenanalysen die Klasse. Eine theoretisch mögliche Analyse in Modellen mit drei Ebenen – Schüler in Klassen in Schulen – ist in DESI aufgrund der geringen Zahl von Schulklassen (maximal zwei) innerhalb der Schulen in der Regel wenig sinnvoll. Modelle mit drei Ebenen wurden daher nur zur Bestimmung der Varianzanteile auf Schul-, Klassen- und Individualebene (vgl. Kapitel 34), nicht jedoch bei Zusammenhangsanalysen verwendet. Die Mehrebenenanalysen in DESI wurden mit dem Programm Mplus (Muthén/Muthén 1998-2007) in der Version 4.0 durchgeführt. Einzelne Analysen, insbesondere die Schätzung von Varianzverteilungen zwischen drei Ebenen (Schüler, Klasse, Schule), wurden mit dem Programm HLM in der Version 6.02 vorgenommen. Diese Software erlaubt die Analyse von Plausible Values und Daten aus Multiplen Imputationen und die Berücksichtigung dieser Datenlage

bei der Schätzung von Standardfehlern und Signifikanzen. Die Berücksichtigung von Gewichten zum Ausgleich unterschiedlicher Teilnahmewahrscheinlichkeiten der Schülerinnen und Schüler (vgl. Kapitel 2) ist mit jedem der verwendeten Analyseprogramme möglich. Soweit nicht explizit anders angegeben, wurden alle in diesem Band berichteten Analysen unter Verwendung der in Kapitel 2 beschriebenen Gewichtung vorgenommen.

Während die Clusterstruktur der Stichprobe bei Analysen von Schülerdaten durch geeignete Methoden berücksichtigt werden muss, besteht dieses Problem bei Daten auf Schulebene nicht und auf Klassenebene in vernachlässigbarem Umfang. Die Schulen, an denen die Daten für DESI erhoben wurden, sind eine echte Zufallsstichprobe (vgl. Kapitel 2), und mit je zwei Klassen pro Schule ist die Clusterung der Stichprobe auf Klassenebene sehr gering. Bei Analysen auf Schul- und Klassenebene (z.B. ausschließlich mit Lehrer- oder Unterrichtsvariablen) kamen daher bei Signifikanztests auch Standardverfahren zum Einsatz.

Bestimmung von Effektgrößen

Die statistische Signifikanz eines Zusammenhangs oder Gruppenunterschieds bedeutet noch nicht zwingend, dass ein gefundener Effekt auch praktisch relevant ist, da bei großen Stichproben auch sehr schwache Effekte signifikant werden können. Zur Frage nach der Signifikanz eines Effektes kommt daher die Frage nach der *Effektgröße* hinzu. Zur Bestimmung der Effektgröße existieren für verschiedene Verfahren der Zusammenhangsanalyse verschiedene Kennwerte. Allen ist gemeinsam, dass untersuchte Effekte unabhängig von den ursprünglichen Skalen der untersuchten Variablen und in über verschiedene Analysen hinweg vergleichbaren Maßeinheiten beschrieben werden sollen. Ebenfalls gemeinsam ist den verschiedenen Effektgrößemaßen das Prinzip, einen untersuchten Zusammenhang oder Gruppenunterschied an der Streuung der interessierenden abhängigen Variablen zu relativieren. Die Größe eines Effektes wird also im Verhältnis zu den in der Stichprobe anzutreffenden Unterschieden zwischen den untersuchten Fällen interpretiert. Dieses Vorgehen lässt sich leicht an einem einfachen fiktiven Beispiel veranschaulichen. Wenn in einer Stichprobe ein Unterschied von zehn Testwert-Punkten zwischen Mädchen und Jungen gefunden wird, ist diese Zahl für sich genommen noch nicht aussagekräftig. Die Einschätzung der praktischen Bedeutsamkeit dieses Geschlechtereffekts, hängt von der Streuung der Testwerte ab. Wenn die Testwerte beispielsweise eine Streuung von $SD = 250$ haben, sind zehn Punkte eine vernachlässigbar kleine Differenz, nämlich nur 4% einer Standardabweichung. Hätten die Testwerte jedoch eine Streuung von z.B. nur 20 Punkten, wären zehn Punkte immerhin eine halbe Standardabweichung – die Geschlechtsdifferenz wäre in diesem Fall deutlich bedeutsamer.

Um die Größe von Effekten sowohl über verschiedene Studien als auch über verschiedene Analysemethoden hinweg vergleichen zu können, wird bei der Darstellung der Ergebnisse auf die übliche von Cohen (1988, 1992) vorgeschlagene

Klassifikation in kleine, mittlere und große Effekte zurückgegriffen. In Tabelle 3.1 sind die in den verschiedenen Ergebnisberichten verwendeten Effektgrößen sowie ihre Klassifikationen aufgelistet (vgl. Cohen 1988, 1992).

Tabelle 3.1: Definition und Klassifikation der bei der Ergebnisdarstellung verwendeten Effektgrößen.

Analyse	Effektgröße	Definition	Klassifikation		
			klein	mittel	groß
Gruppenunterschiede	d	An der Streuung standardisierte Differenz $d = \frac{M_A - M_B}{\sigma}$	0.20	0.50	0.80
Bivariate Zusammenhänge	r	An den Streuungen relativierte Kovarianz $r = \frac{\text{cov}(x, y)}{\sigma(x) \cdot \sigma(y)}$	0.10	0.30	0.50
Multiple Regression	R ²	An der Gesamtvarianz relativierte durch alle Prädiktoren erklärte Varianz $R^2 = \frac{QS_{\text{Modell}}}{QS_{\text{Gesamt}}}$	0.02	0.13	0.26
Varianzanalysen	partielles η^2	An der Fehlervarianz relativierte Quadratsumme des Effekts $\eta^2 = \frac{QS_{\text{Effekt}}}{QS_{\text{Effekt}} + QS_{\text{Fehler}}}$	0.01	0.06	0.14

Anmerkung: QS = Quadratsumme

In Mehrebenenmodellen ist der Anteil der erklärten Varianz nicht so einfach zu bestimmen und zu interpretieren wie bei Analysen ohne Berücksichtigung der Mehrebenenstruktur (vgl. Hox 2003; Snijders/Bosker 1994). Bei den in diesem Band berichteten Ergebnissen aus Mehrebenenanalysen werden daher keine Effektgrößen angegeben, sondern für die einzelnen Effekte innerhalb der Modelle standardisierte Koeffizienten berichtet. Pfad- oder Regressionskoeffizienten werden hierfür an den Streuungen der jeweiligen unabhängigen und abhängigen Variablen standardisiert (vgl. Hox 2003), so dass sie in einem Wertebereich von -1 bis 1 liegen und von der Größenordnung entsprechend Korrelationskoeffizienten interpretiert werden können.

Modelle mit latenten Variablen

Bei der empirischen Erhebung von Schüler-, Klassen- und Schulmerkmalen werden häufig mehrere Indikatoren erfasst, welche auf dasselbe theoretische Konstrukt abzielen. Ein geläufiges Beispiel hierfür sind Fragebogenskalen, in denen dasselbe

Konstrukt (z.B. Leistungsmotivation) mit mehreren Fragen erfasst wird. Die Messung eines theoretischen Konstrukts mit Hilfe mehrerer verschiedener Indikatoren ist eine weit verbreitete Methode und dient vor allem einer höheren Messgenauigkeit (Reliabilität) und einer breiteren Generalisierbarkeit (Validität) der resultierenden Werte. Oft werden diese Messwerte durch Summen- oder Mittelwertsbildung aus den einzelnen Indikatoren gebildet, dieses Vorgehen wurde in DESI für die Fragebogendaten gewählt.

Es ist jedoch auch möglich, die Idee eines theoretischen Konstrukts, das mehreren beobachteten Variablen zugrunde liegt, direkt in statistische Analysemodelle einzubeziehen. In solchen Modellen werden die theoretischen Konstrukte als *latente Variablen* betrachtet, auf die über die gemessenen Indikatoren geschlossen werden kann. Modelle mit latenten Variablen erlauben eine messfehlerbereinigte Schätzung von Zusammenhängen zwischen den interessierenden Konstrukten. Auch die zur Analyse der Leistungstests verwendeten IRT-Modelle (s.o.) sind Modelle mit latenten Variablen, wobei die einzelnen Testaufgaben die beobachteten Indikatoren darstellen. Soweit in den Analysen zu DESI mit PVs gerechnet wird, werden die Kompetenzen statistisch also schon als latente Variablen behandelt. In vielen Analysen auf Basis von Fragebogendaten ist es ebenfalls so, dass mehrere beobachtete Variablen auf ein übergeordnetes Konstrukt zurückgeführt werden, z.B. verschiedene konkrete Formen der Kooperation in Fachkollegien auf die allgemeine Kooperativität an den jeweiligen Schulen (vgl. Kapitel 34). In diesen Fällen wurden Modelle mit latenten Variablen angewandt, in denen beobachtete numerische Werte (z.B. aus Fragebogenskalen) als Indikatoren für latente Variablen verwendet und die Zusammenhänge der interessierenden Konstrukte auf latenter Ebene untersucht wurden. Diese Modelle mit quantitativen Indikatoren werden in der Literatur häufig unter dem Begriff *Strukturgleichungsmodelle* beschrieben (z.B. Bollen 1989). Wenn die latenten Variablen nicht kontinuierlich, sondern kategorial sind, wird von *latenten Klassenanalysen* (z.B. Rost 2004) gesprochen. Die beobachteten Variablen werden in diesem Fall als Indikatoren für die Zugehörigkeit zu bestimmten Gruppen (latenten Klassen) betrachtet, wobei die Gruppenzugehörigkeit eine nicht latente – also nicht direkt beobachtbare – Variable darstellt.

Auch die Analysen mit latenten Variablen wurden in DESI mit Mplus 4.0 durchgeführt. Es ist möglich, Modelle mit latenten Variablen mit Mehrebenenmodellen zur Berücksichtigung der Stichprobenstruktur (s.o.) zu kombinieren, hieraus resultieren z.B. Mehrebenen-Strukturgleichungsmodelle, die in den Analysen zu Unterrichts- und Schulvariablen zum Einsatz kamen.

Bei der Analyse von Modellen mit latenten Variablen ist es wichtig zu beachten, wie gut das verwendete Modell mit den tatsächlich beobachteten Daten übereinstimmt. Um die Güte dieser Übereinstimmung zu beurteilen, können verschiedene statistische Indizes für die Modellanpassung (*model fit*) herangezogen werden. Diese Gütemaße basieren alle auf der Übereinstimmung der beobachteten Daten mit dem Datenmuster, dass bei Gültigkeit des geschätzten Modells zu erwarten wäre. Bei einigen dieser Gütemaße wie dem *Goodness of Fit Index* (GFI) sind möglichst hohe

Werte nahe eins wünschenswert, bei anderen, wie dem *Root Mean Square Error of Approximation* (RMSEA) möglichst niedrige Werte nahe Null. Tabelle 3.2 gibt einen Überblick darüber, welche Werte für verschiedene in diesem Band verwendete Gütemaße als Hinweise auf eine gute Modellanpassung interpretiert werden. Eine detaillierte Beschreibung verschiedener Gütemaße findet sich bei Schermelleh-Engel und Moosbrugger (2002).

Tabelle 3.2: Beurteilung der Güte von Modellen mit latenten Variablen anhand ausgewählter Gütekriterien.

Gütekriterium	Akzeptable Modellanpassung	Gute Modellanpassung
χ^2 / df (Chi-Quadrat-Prüfstatistik geteilt durch Freiheitsgrade)	≤ 3.00	≤ 2.00
SRMR (Standardized Root Mean Square Residual)	≤ 0.08	≤ 0.05
RMSEA (Root Mean Square Error of Approximation)	≤ 0.08	≤ 0.05
GFI (Goodness of Fit Index)	≥ 0.90	≥ 0.95
NNFI (Non-Normed Fit Index)	≥ 0.90	≥ 0.95
CFI (Comparative Fit Index)	≥ 0.90	≥ 0.95

Verwendung von Kontrollvariablen

Bei der Untersuchung der Zusammenhänge der Kompetenzen mit anderen Variablen ist zu berücksichtigen, dass die in DESI erhobenen Größen in vielfältiger Weise miteinander korreliert sind. Wenn Unterschiede in den sprachlichen Kompetenzen mit einer bestimmten Variablen erklärt werden sollen, muss immer berücksichtigt werden, dass diese Variable mit anderen möglichen erklärenden Größen *konfundiert* ist. Wenn z.B. der sozioökonomische Status der Eltern mit den kognitiven Grundfähigkeiten der Kinder zusammenhängt, muss dieser Abhängigkeit Rechnung getragen werden, wenn der spezifische Effekt des sozioökonomischen Status auf Unterschiede in sprachlichen Kompetenzen beurteilt werden soll. Hierzu kann der Effekt unterschiedlicher kognitiver Fähigkeiten statistisch kontrolliert werden, indem diese Variable z.B. in einer Regressionsanalyse zusammen mit dem sozioökonomischen Status als Prädiktor verwendet wird. In diesem Fall würde die kognitive Grundfähigkeit als *Kontrollvariable* verwendet. Der Effekt des sozioökonomischen Status würde dann als der Effekt interpretiert, den diese Variable hätte, wenn alle Schüler hinsichtlich ihrer kognitiven Grundfähigkeit gleich wären.

Welche Variable in einem Modell eine „Kontrollvariable“ darstellt, ist eine jeweils aufgrund inhaltlicher Überlegungen zu entscheidende Frage. Der Einbezug von Kontrollvariablen ist in DESI insbesondere dort wichtig, wo schul- und unterrichtsbezogene Variablen mit Schülermerkmalen konfundiert sind, auf welche die Schule keinen Einfluss hat. Die Idee dieser Kontrolle lässt sich anschaulich ausdrücken, indem man z.B. positive Effekte bestimmter Unterrichtstechniken dahingehend

beschreibt, dass diese Praktiken mit *erwartungswidrig* guten Kompetenzen einhergehen. Damit ist gemeint, dass die Kompetenzen der Schüler in Klassen, in denen diese Unterrichtspraxis verstärkt zum Einsatz kam, höher sind, als z.B. aufgrund ihrer kognitiven Grundfähigkeiten und ihres Elternhauses zu erwarten gewesen wäre. Diese Abweichungen von den aufgrund der individuellen Ausgangsvoraussetzungen zu erwartenden Kompetenzen werden in der Schulforschung auch als *value added* bezeichnet, d.h. als der zusätzliche Beitrag zur Kompetenzentwicklung, der über außerhalb der Schule liegende Einflussfaktoren (z.B. die soziale Herkunft der Schülerinnen und Schüler) hinaus geht.

Neben den Schülermerkmalen werden in vielen Analysen zu DESI auch Unterschiede zwischen den verschiedenen Bildungsgängen kontrolliert. Dies ist wichtig, wenn z.B. bestimmte Unterrichtsmerkmale in Hauptschulen häufiger zu beobachten sind, und die Wirkung dieser Merkmale unabhängig davon eingeschätzt werden soll, dass in Hauptschulklassen insgesamt niedrigere Kompetenzen zu beobachten sind. Auch bei den auf individueller Ebene erfassten Kontrollvariablen kann die Variation auf verschiedenen Analyseebenen berücksichtigt werden. Variablen wie die kognitiven Grundfähigkeiten oder der soziale Status können einerseits auf individueller Ebene berücksichtigt werden, andererseits aber auch auf Schul- oder Klassenebene. Auf höherer Ebene wird dann die Zusammensetzung der Schüler in einer Klasse oder Schule kontrolliert, die entsprechenden Kontrollvariablen können hierzu durch Aggregation der individuellen Werte gewonnen werden.

Zu den in den Analysen der DESI-Daten verwendeten Kontrollvariablen gehören, soweit nicht anders angegeben, durchweg der Bildungsgang, der individuelle Sprachhintergrund (Erstsprache) und der Anteil der Schüler mit nicht deutscher Erstsprache, die kognitiven Grundfähigkeiten des einzelnen Schülers und der einzelnen Schülerin sowie der entsprechende Klassendurchschnitt, das Geschlecht sowie der Mädchenanteil in der Klasse, der sozioökonomische Status der Herkunftsfamilie und die soziale Komposition der Klasse. Der Bildungsgang wurde bei der Verwendung als Kontrollvariable in der Regel in zwei *Dummy-Variablen* kodiert, in denen zum einen die Zugehörigkeit zum Gymnasium, zum anderen die Zugehörigkeit zur Realschule mit eins, die jeweils anderen Bildungsgänge mit null kodiert wurden. Hierdurch wurden Hauptschule und IGS gemeinsam als Referenzkategorie verwendet.

Es ist zu beachten, dass die in DESI erhobenen Daten auch unter Berücksichtigung von vielfältigen Kontrollvariablen in praktisch keinem Fall eine sichere *kausale* Aussage darüber zulassen, auf welche Variablen beobachtete Kompetenzunterschiede ursächlich zurückgeführt werden können. Dennoch ist die Verwendung von Kontrollvariablen wichtig, um den wechselseitigen Abhängigkeiten und relativen Bedeutsamkeiten der verschiedenen erklärenden Variablen Rechnung zu tragen.

Vorhersage von Kompetenzzuwächsen

Durch die in DESI vorgenommene Messung zu zwei Zeitpunkten kann untersucht werden, welche Variablen mit Unterschieden in Kompetenzzuwächsen innerhalb

des neunten Schuljahres zusammenhängen. Diese Variablen können sowohl individueller Natur sein (gibt es bestimmte förderliche Bedingungen im Elternhaus), aber auch auf Klassenebene zu suchen sein (welche Unterrichtsmerkmale gehen mit einem hohen Kompetenzzuwachs einher) oder auf Schulebene liegen (welche Charakteristika der Fachkollegien gehen mit höheren Kompetenzzuwächsen einher). Die Vorhersage von Veränderungswerten – d.h. von Differenzen zwischen zwei Zeitpunkten – ist in empirischen Untersuchungen generell mit Schwierigkeiten verbunden, da diese Werte höhere Messfehler beinhalten und, bedingt durch Messfehler und durch zeitpunktspezifische Variation, häufig negativ mit dem Ausgangswert korreliert sind (z.B. Rost 2004). Angesichts dieser Problematik ist die Analyse von Veränderungswerten ein in der sozialwissenschaftlichen Forschung kontrovers diskutiertes Thema (z.B. Campbell/Kenny 2003; Allison 1990). Aufgrund der genannten problematischen Eigenschaften von Veränderungswerten werden diese häufig nicht als abhängige Variablen verwendet. Stattdessen werden die Werte zum Ende der neunten Jahrgangsstufe als abhängige Variable und die Werte zu Beginn der neunten Jahrgangsstufe als Prädiktor in das Modell einbezogen, um interindividuelle Unterschiede in den Ausgangswerten statistisch zu kontrollieren. Dieses sogenannte kovarianzanalytische Vorgehen führt jedoch unter bestimmten Bedingungen, insbesondere wenn der Ausgangswert mit einem interessierenden Prädiktor korreliert ist, zur Schätzung artifizierlicher Effekte – es resultieren Zusammenhänge von Drittvariablen mit den „kontrollierten“ Messwerten zum zweiten Zeitpunkt, obwohl in Wahrheit keine Zusammenhänge mit den Zuwächsen zwischen den Zeitpunkten bestehen. Die Verwendung von Differenzwerten als abhängige Variable führt unter den meisten Bedingungen hingegen zu korrekten Schätzungen der interessierenden Effekte (Allison 1990; Hartig/Jude 2006, Jude/Hartig 2006). In DESI wurden daher, wenn der Kompetenzzuwachs von Interesse war, die PV-Differenzen für die zu beiden Messzeitpunkten in der neunten Jahrgangsstufe eingesetzten Tests als abhängige Variablen verwendet.

Ausnahmen bilden jene Analysen, in denen beide Messzeitpunkte in der neunten Jahrgangsstufe gezielt aufgenommen wurden, um zum Beispiel Effekte des Unterrichts in Abhängigkeit von der Leistung zum ersten Messzeitpunkt zu berechnen. In Mehrebenen-Pfadmodellen wurde in diesem Fall die Leistung zu Beginn und zum Ende der neunten Jahrgangsstufe in ihrem Zusammenhang mit Unterrichtsvariablen modelliert, um die Adaptivität des Unterrichts zu untersuchen. In diesen Modellen wurden für jeden der beiden Messzeitpunkte in der neunten Jahrgangsstufe Zusammenhänge zwischen der Leistung und den oben angeführten Hintergrundvariablen modelliert.

Literatur

Adams, R. J./Wilson, M./Wang, W. (1995). The multidimensional random coefficients multinomial logit model. In: *Applied Psychological Measurement*, 21, S. 1-23.

- Adams, R. J./Wu, M. L. (2007). The mixed-coefficients logit model: A generalized form of the Rasch model. In: von Davier, M./Carstensen, C. H. (Hrsg.): *Multivariate and mixture distribution Rasch models*. New York, Berlin: Springer, S. 57-75.
- Adams, R. J./Wu, M. L./Carstensen, C. H. (2007). Application of multivariate Rasch models in international large-scale educational assessments. In: von Davier, M./Carstensen, C. H. (Hrsg.): *Multivariate and mixture distribution Rasch models*. New York, Berlin: Springer, S. 271-280.
- Allison, P. D. (1990). Change scores as dependent variables in regression analysis. In: *Sociological Methodology*, 20, S. 93-114.
- Beaton, E./Allen, N. (1992). Interpreting scales through scale anchoring. In: *Journal of Educational Statistics*, 17, S. 191-204.
- Beck, B./Klieme, E. (2007). *Sprachliche Kompetenzen – Konzepte und Messung*. Weinheim: Beltz.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Brick, M./Morganstein, D./Valliant, R. (2000). *Analysis of Complex Sample Data Using Replication*. Rockville: Westat.
- Campbell, D. T./Kenny, D. A. (2003). *A primer on regression artifacts*. New York, London: Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. In: *Psychological Bulletin*, 112, S. 155-159.
- Fay, R. E. (1989). Theory and application of replicate weighting for variance calculations. In: *Proceedings of the Section on Survey Research of the American Statistical Association*, S. 212-217.
- Ganzeboom, H. B. G./de Graaf, P. M./Treiman, D. J./de Leeuw, J. (1992): A standard international socio-economic index of occupational status. In: *Social Science Research* 21, S. 1-56.
- Goldhammer, F./Hartig, J. (2007). Interpretation von Testresultaten und Testeichung. In: Moosbrugger, H./Kelava, A. (Hrsg.): *Test- und Fragebogenkonstruktion*. Berlin: Springer, S. 165-192.
- Harsch, C./Schröder, K. (2007). Textrekonstruktion: C-Test. In: Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen – Konzepte und Messung*. Weinheim: Beltz, S. 212-225.
- Hartig (2007). Skalierung und Kompetenzniveaus. In Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen – Konzepte und Messung*. Weinheim: Beltz, S. 83-99.
- Hartig, J./Frey, A. (2006). Using plausible values from multidimensional irt models to estimate change in large scale assessments. Paper presented at the 15th international meeting of the Psychometric Society in Montréal, June 14–17 2006.
- Hartig, J./Jude, N. (2005). Effects of different estimators for student proficiencies on multidimensional multilevel structures. Paper presented at the Fifth International Amsterdam conference on Multilevel Analysis, Amsterdam, March 21-23 2005.
- Hartig, J./Jude, N. (2006). Einschätzung von Unterrichtswirksamkeit auf Basis von (nur) zwei Messzeitpunkten. Vortrag auf der 68. Tagung der Sektion Empirische Bildungsforschung der DGfE in München vom 10. bis 13. September 2006.
- Hartig, J./Kühnbach, O. (2006). Schätzung von Veränderung mit Plausible Values in mehrdimensionalen Rasch-Modellen. In: Ittel, A./Merkens, H. (Hrsg.): *Veränderungsmessung und Längsschnittstudien in der Erziehungswissenschaft*. Wiesbaden: Verlag für Sozialwissenschaften, S. 27-44.
- Heller, K. A./Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen. Revision*. Göttingen: Beltz.
- Helmke, A./Hosenfeld, I. (2004): Vergleichsarbeiten – Standards – Kompetenzstufen: Begriffliche Klärungen und Perspektiven. In: Jäger, R.S./Frey, A. (Hrsg.): *Lernprozesse, Lernumgebung und Lerndiagnostik. Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert*. Landau: Verlag Empirische Pädagogik.

- Hox, J. (2003). *Multilevel analysis: Techniques and applications*. Mahwah: Lawrence Erlbaum.
- Jude, N./Hartig, J. (2006). Modelle zur Vorhersage von Leistungszuwächsen bei der Untersuchung von Unterrichtseffekten. Vortrag auf dem 45. Kongress der Deutschen Gesellschaft für Psychologie in Nürnberg vom 17. bis 21. September 2006.
- Klieme, E./Baumert, J./Köller, O./Bos, W. (2000). Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung von Kompetenzen. In: Baumert, J./Bos, W./Lehmann, R. (Hrsg.): TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie. Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen: Leske + Buderich, S. 85-133.
- Meiser, T./Stern, E./Langeheine, R. (1998). Latent Change in Discrete Data: Unidimensional, Multidimensional, and Mixture Distribution Rasch Models for the Analysis of Repeated Observations. In: *Methods of Psychological Research Online*, 3, S. 75-93.
- Mislevy, R. J./Beaton, A. E./Kaplan, B./Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of responses. In: *Journal of Educational Measurement*, 29, S. 133-161.
- Muthén, L.K./Muthén, B.O. (1998-2007). *Mplus User's Guide*. Fourth Edition. Los Angeles, CA: Muthén & Muthén
- Nold, G./Rossa, H. (2007). Sprachbewusstheit. In: Beck, B./Klieme, H. (Hrsg.): *Sprachliche Kompetenzen – Konzepte und Messung*. Weinheim: Beltz, S. 226-244.
- OECD (2001). *Knowledge and Skills for Life. First Results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.
- OECD (2004). *Learning for Tomorrow's World – First Results from PISA 2003*. Paris: OECD.
- Raghunathan T. E./Solenberger P. W./Hoewyk J. V. (2002). *IVEware: Imputation and Variance Estimation Software. Installation Instructions and User Guide*. Survey Research Center, Institute for Social Research, University of Michigan.
- Rauch, D./Hartig, J. (2007). Interpretation von Testwerten in der IRT. In Moosbrugger, H./Kelava, A. (Hrsg.): *Test- und Fragebogenkonstruktion*. Berlin: Springer, S. 240-250.
- Raudenbush, S./Bryk, A./Congdon, R. (2004). *HLM 6 for Windows*. Chicago, IL: Scientific Software.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. 2. überarbeitete und erweiterte Auflage. Bern, Göttingen: Huber.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.
- Schafer, J. L./Olsen, M. (1998). Multiple imputation for multivariate missing-data problems: a data analyst's perspective. In: *Multivariate Behavioral Research* 33, S. 545-571.
- Schermelleh-Engel, K./Moosbrugger, H. (2002). Beurteilung der Modellgüte von Strukturgleichungsmodellen. Frankfurt: Arbeiten aus dem Institut für Psychologie der J. W. Goethe-Universität, Heft 4/2002.
- Snijders, T.A.B./Bosker, R.J. (1994). Modeled variance in two-level models. In: *Sociological Methods and Research*, 22, S. 342-363.
- Walter, O. (2006). *Kompetenzmessung in den PISA-Studien. Simulationen zur Schätzung von Verteilungsparametern und Reliabilitäten*. Lengerich: Pabst.
- Westat (2000). *WesVar 4.0 User's guide*. Rockville: Westat.
- Wilson, M. R. (2003). On choosing a model for measuring. *Methods of Psychological Research Online*, 8, 1-22.
- Wolf, A. (2006). *Shorter Tests Through the Adaptive Use of Planned Missing Data in Sampling Designs*. Unpublished PhD Thesis, Friedrich-Schiller University Jena.
- Wu, M. L./Adams, R. J./Wilson, M. R. (1998). *ConQuest: Generalized item response modelling software*. Melbourne: Australian Council for Educational Research.

Hans-Günter Rolff / Jan von der Gathen

4 Rückmeldungen an Lehrkräfte und Rezeption

Rückmelde- und Rezeptionsstudien sind inzwischen Bestandteil aller groß angelegten Schülerleistungsstudien sowohl bei Stichprobenstudien wie PISA (Watermann/Stanat 2004), IGLU (Schwippert 2004) oder QuaSUM (Peek 2004) als auch bei Vollerhebungen wie LAU (Klug/Reh 2000) oder MARKUS (Schrader/Helmke 2004).

Übergreifend gilt nach den Befunden dieser Studien, dass die erfolgreiche inner-schulische Rezeption und Umsetzung von Evaluationsergebnissen von einer Vielzahl von Bedingungen abhängen. Hierzu gehören die Gestaltung der Rückmeldungen, das Vorwissen, die Einstellungen und die Handlungskompetenzen der Lehrkräfte in den Schulen und schließlich Unterstützungsleistungen der eine Schule begleitenden Institutionen, insbesondere der Schulaufsicht und der Fortbildungseinrichtungen. Insgesamt zeigte sich, dass die Rezeption durch Lehrkräfte sehr schwach ausgeprägt ist und der Rücklauf der Lehrerfragebögen zur Rezeption der Befunde weit unter den üblichen Quoten bei Lehrerbefragungen liegt. Dabei ist die Rezeptionsneigung bei Stichprobenerhebungen offenbar noch geringer als bei Vollerhebungen, was vermutlich auch daran liegt, dass Stichprobenerhebungen gar nicht für Zwecke der Schulentwicklung konzipiert sind: International vergleichende Leistungsstudien sind vornehmlich für die Systemebene gedacht und dafür auch nützlich (vgl. Baumert 2001). Stichprobe und Design solcher Studien sind konzipiert, um auf nationaler Ebene repräsentative Ergebnisse zu generieren. Dennoch kann die Einzelschule nicht als Adressat einer Ergebnissrückmeldung im Rahmen von Leistungsvergleichsstudien ausgespart werden; denn es entspricht der Forschungsethik und erhöht auch die Teilnahmebereitschaft, wenn Resultate an diejenigen Institutionen und Personen rückgemeldet werden, bei denen sie erhoben wurden.

4.1 Anlage der Rückmeldungen und der Rezeptionsstudie

Wie in Kapitel 2 begründet, wurde die DESI-Studie nicht mit vergleichender Absicht konzipiert und durchgeführt. Sie verfolgte auch nicht das Ziel, Entwicklungsprozesse in einzelnen Schulen oder gar Klassen anzuregen. Weil aber insgesamt bisher wenig über die innerschulische Rezeption und Nutzung der rückgemeldeten Ergebnisse von Schulleistungsstudien bekannt ist, wurde trotz dieses Umstandes eine Rezeptionsstudie durchgeführt. Sie umfasste neun Fallstudien in drei Bundesländern, wobei jede Schulart vertreten war, sowie einen *survey* zur Befragung der Fachlehrkräfte Deutsch und Englisch.

Grundlage sowohl der ausführlichen Fallstudien als auch des *survey* waren Rückmeldungen ausgewählter Ergebnisse, die interessierten Lehrerinnen und Lehrern im Mai/Juni 2005 zur Verfügung gestellt wurden. Diese Rückmeldungen basierten auf einem durch die Amtschefkommission der KMK genehmigten und den Schulleitungen vorab übermittelten Konzept. Gemäß der vereinbarten Vorgehensweise bestanden die Rückmeldungen aus zwei Teilen: Einem allgemeinen Teil, der die DESI-Studie noch einmal kurz vorstellte, Teilnahmequoten erläuterte und das Prinzip der Rückmeldungen erklärte – diesen Teil konnten auch die Schulleitungen beantragen. Der zweite Teil richtete sich ausschließlich an die Lehrkräfte für Deutsch und Englisch. In ihm wurden ausgewählte konkrete unterrichts- und klassenbezogene Ergebnisse mitgeteilt.

Durch die DESI-Rückmeldungen sollten die beteiligten Lehrkräfte für Deutsch und Englisch die Möglichkeit erhalten, sach- und ergebnisbezogen über den eigenen Unterricht zu reflektieren und gegebenenfalls Optimierungsansätze für ihren Lehr-Lern-Prozess zu finden. Dementsprechend waren also diejenigen Fachlehrer für Deutsch und Englisch die Adressaten der klassen- und unterrichtsbezogenen Rückmeldungen, welche die Rückmeldung im IEA DPC angefordert hatten. Es wurde ausschließlich der Stand ausgewählter sprachlicher Leistungen am Ende der neunten Jahrgangsstufe mitgeteilt. Die Anforderungen der Tests wurden durch Itembeispiele illustriert und die Testleistungen anhand grafischer Darstellungen der Leistungsverteilungen in der zutreffenden Klasse rückgemeldet. Zusätzlich zu der tatsächlich beobachteten Verteilung wurde für jede Klasse und jeden Test eine Vergleichsverteilung angezeigt, die verdeutlichte, welche Verteilung auf Basis der gesamten DESI-Stichprobe für eine Klasse desselben Bildungsgangs und unter Beachtung bestimmter Schülervariablen zu erwarten gewesen wäre. Diese Verteilung wurde berechnet¹, um den Lehrkräften eine Einordnung der tatsächlichen Klassenverteilung und damit eine interpretative Orientierung zu ermöglichen.

Am Ende der neunten Jahrgangsstufe wurden die Schülerinnen und Schüler unter anderem auch zum Unterricht in ihrer Klasse bzw. in ihrem Kurs befragt (Unterrichtswahrnehmung durch den Schüler). Auch hierzu wurden einige Rückmeldungen gegeben, und zwar zur Klassenführung und zur Motivierung der Schülerinnen und Schüler für die fachlichen Inhalte. Schließlich wurden noch ausgewählte Ergebnisse der DESI-Elternbefragung mitgeteilt.

Zwischen der Datenerhebung am Ende der neunten Jahrgangsstufe und den beschriebenen Rückmeldungen ausgewählter Ergebnisse an die Deutsch- und Englischlehrkräfte lag ein Zeitraum von ca. einem Jahr.

Eine Ausnahme wurde lediglich bei den neun Fallstudien-Schulen gemacht. Hier erhielten die Fachlehrkräfte Deutsch/Englisch bereits nach dem ersten Messzeitpunkt zu Beginn der neunten Jahrgangsstufe eine erste kurze Rückmeldung. Neben einem

1 Die Vergleichsverteilungen basierten auf Regressionsanalysen, in denen Bildungsgang, Geschlecht, sozioökonomischer Status und kognitive Grundfähigkeiten zur Vorhersage der Schülerleistungen verwendet wurden. Zur Erzeugung der Vergleichsverteilungen wurde auch der Vorhersagefehler dieser Modelle berücksichtigt.

einführenden Textteil wurde für jede dort betrachtete Domäne eine Kurzbeschreibung des Testmoduls zur Verfügung gestellt. Darunter war ein Balkendiagramm mit der prozentualen Verteilung der Schülerinnen und Schüler der betreffenden Klasse auf den Kompetenzniveaus dargestellt. Aufgabenbeispiele konnten zu diesem Zeitpunkt noch nicht öffentlich gemacht werden. Adressat war auch hier die Einzellehrkraft; die Leistungsrückmeldungen wurden an deren Privatadresse verschickt. Diese Zustellung erster Ergebnisse an die Lehrkräfte der neun Fallschulen war die Basis für ausführliche Fallstudien an diesen Schulen, die nach den allgemein gegebenen DESI-Rückmeldungen durchgeführt und durch Interviews begleitet wurden.

Parallel zu den ausführlichen Fallstudien wurde eine Lehrkräftebefragung an allen 219 DESI-Schulen durchgeführt (*survey*). Diese erfasste sechs Wochen nach den allgemeinen DESI-Rückmeldungen die Ergebnisse des zweiten Messzeitpunktes die kurzfristig zu beobachtenden Wirkungen und Verarbeitungsstrategien in den Schulen.

4.2 Kernergebnisse der Fallstudien

Nach der inhaltsanalytischen Auswertung des Interviewmaterials lassen sich folgende Aussagen der Lehrkräfte zusammenfassen:

- In den Fallschulen standen alle befragten Lehrkräfte externer Leistungsmessung offen gegenüber. Einige Lehrkräfte wünschen sich eine wesentlich größere Transparenz und Aufklärung über die konkrete inhaltliche Testgestaltung anhand von Aufgabenbeispielen.
- Der Faktor „Zeit“ stellte sich als besonders wichtig heraus. Sämtliche Lehrkräfte beklagten den mit zwölf Monaten aus ihrer Sicht viel zu langen Zeitraum zwischen Datenerhebung und detaillierter Ergebnisrückmeldung.
- Insgesamt wurde das vorliegende Rückmeldeformat als überwiegend verständlich eingeschätzt.
- Eine kooperative Verarbeitung der Rückmeldungen hat nur in solchen Kollegien stattgefunden, in denen bereits eine Kultur der offenen Kommunikation und Kooperation gepflegt wurde.
- Viele Lehrkräfte stellten beim Umgang mit den DESI-Rückmeldungen immer wieder auch den direkten Bezug zum Einzelschüler her. Einen solchen individuellen Bezug können und wollen Rückmeldungen aus *large scale assessments* jedoch nicht leisten. Das war den meisten Lehrkräften nicht klar.

4.3 Kernergebnisse des Surveys

Allen beteiligten Lehrkräften aus den 219 Schulen wurde das Angebot unterbreitet, eine Rückmeldung über ausgewählte Ergebnisse des zweiten Testzeitpunkts zu erhalten. Von 774 in Frage kommenden Lehrkräften forderten 137 Lehrerinnen und Lehrer eine klassenbezogene Auswertung beim IEA DPC an. An diesen Personenkreis wurde der *survey*-Fragebogen sechs Wochen nach Versenden der DESI-Rückmeldungen

gerichtet. Der bei diesen Befragten erzielte Rücklauf von 36% ist zwar nicht zufrieden stellend, aber größer als bei anderen Rezeptionsstudien im Rahmen von *large scale assessments*. Von den ausgefüllten Fragebögen kommen 47% aus Gymnasien, 25% aus Realschulen, 16% aus Schulen mit mehreren Bildungsgängen und nur jeweils 4% aus Haupt- bzw. Gesamtschulen. Mit 63% des Gesamtrücklaufs haben die Englisch-Fachlehrkräfte deutlich mehr Interesse an der Rezeptionsstudie gezeigt als die Deutsch-Fachlehrkräfte.

Es sollte bei den folgenden Ausführungen bewusst sein, dass die Antwortenden eine „Positivauswahl“ darstellen, denn eine Lehrkraft, die keinerlei Interesse an der DESI-Studie und an den Ergebnismeldungen gehabt hat, dürfte kaum den Fragebogen ausgefüllt haben. Die antwortenden Lehrkräfte aber haben sich zu zwei Dritteln intensiv und gründlich mit den Datenrückmeldungen auseinandergesetzt. Dabei betrug die durchschnittliche Zeitinvestition in die Informationsentnahme (Lesen und Interpretieren) drei Stunden.

Es blieb den Lehrkräften überlassen, die klassenbezogenen Ergebnisse weiteren Lehrerinnen und Lehrern, die in der jeweiligen Klasse ebenfalls Unterricht erteilten, zugänglich zu machen. Eine untergeordnete Rolle spielten in einer „kooperativen Auseinandersetzung“ andere Deutsch- oder Englisch-Lehrkräfte und die Gesamtkonferenz. In den Fachkonferenzen Deutsch und Englisch wurden die Rückmeldungen zwar angesprochen, jedoch nicht eingehender behandelt (36% bzw. 40% haben sich „halbwegs“ mit der Rückmeldung beschäftigt).

Fast alle Lehrkräfte beurteilten die Verständlichkeit der Rückmeldungen auf einer vierstufigen Skala („gut verständlich“ bis „nicht verständlich“) positiv. Die Beschreibung der Einzeltests wird von 49% als „gut verständlich“ und 47% als „eher verständlich“ eingeschätzt. Ähnlich verhält es sich mit der Beurteilung der gewählten Darstellungsweise: Die Darstellungsform der Testergebnisse – Boxplots – finden 45% „gut verständlich“ und 46% „eher verständlich“; die Balkendiagramme (Darstellungsform der Befragungsergebnisse) werden zu 70% als „gut“ und zu 29% als „eher verständlich“ eingestuft und die Beispielaufgaben zu 65% bzw. 33% als gut lesbar beurteilt.

Neben der Verständlichkeit der Rückmeldungen spielt deren eingeschätzte Nützlichkeit für das konkrete Unterrichtshandeln eine große Rolle. Insgesamt stimmte ein Fünftel der Lehrkräfte zu, dass Rückmeldungen für die weitere Arbeit prinzipiell sehr nützlich sein können. Allerdings gaben auch 98% der Deutsch- und 93% der Englisch-Lehrkräfte an, im Ergebnis der Rückmeldungen keine Änderungen im Fachunterricht in den betreffenden Domänen vorgenommen zu haben. Ein solches Ergebnis war zu erwarten (vgl. Bosen/von der Gathen 2004). Besser als groß angelegte Stichprobenuntersuchungen sind dafür offenbar Vollerhebungen mit zeitnahen Rückmeldungen geeignet (vgl. Müller 2006), deren Testaufgaben noch lehrplannäher angelegt werden können.

Literatur

- Baumert, J. (2001): Evaluationsmaßnahmen im Bildungsbereich. In: Newsletter – Österreichische Gesellschaft für Forschung und Entwicklung im Bildungswesen 2, H. 1, S. 3-15.
- Bonsen, M./von der Gathen, J. (2004): Schulentwicklung und Testdaten – die innerschulische Verarbeitung von Leistungsrückmeldungen. In: Holtappels, H. G./Klemm, K./Pfeiffer, H./Rolf, H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Band 13. Weinheim: Juventa, S. 225-252.
- Klug, C./Reh, S. (2000): Was fangen die Schulen mit den Ergebnissen an? In: Pädagogik 12, H. 52, S. 16-21.
- Müller, S. (2006): Wie Schulen gut mit den Ergebnissen der Vergleichstests arbeiten. In: Buchen, H./Horster, L./Rolf, H.-G. (Hrsg.): Schulleitung und Schulentwicklung. Ein Reader. Berlin: Raabe.
- Peek, R. (2004): Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (Quasum). In: Empirische Pädagogik 18, H. 1, S. 82-114.
- Schrader, W./Helmke, A. (2004): Von der Evaluation zur Innovation? Die Rezeptionsstudie WALZER: Ergebnisse der Lehrerbefragung. In: Empirische Pädagogik 18, H. 1, S. 140-161.
- Schwippert, K. (2004): Leistungsrückmeldungen an Grundschulen im Rahmen der Internationalen Grundschul-Lese-Untersuchung (IGLU). In: Empirische Pädagogik 18, H. 1, S. 62-81.
- Watermann, R./Stanat, P. (2004): Schulrückmeldungen in PISA 2000. In: Empirische Pädagogik 18, H. 1, S. 40-61.

Steffen Gailberger / Heiner Willenberg

5 Leseverstehen Deutsch

Was ist Textverstehen? Ein Angler fängt an der Donau einen großen Fisch. Der sagt: „Wirf mich zurück, so hast du drei Wünsche frei.“ Der Angler antwortet: „Ich möchte ein Schloss haben, reich sein und eine Prinzessin zur Frau.“ „Gemacht“, sagt der Fisch, „die Zeit egal?“ „Die Zeit ist mir egal“, erwidert der Angler. Als dieser am nächsten Morgen erwacht, befindet er sich in einem Schloss, entdeckt edle Juwelen darin und findet auf dem Bett neben sich die Prinzessin liegen. Diese lächelt ihn an und sagt: „Schön dass du endlich wach bist, Franz-Ferdinand? Wir fahren nämlich heute nach Sarajewo.“

An einem solchen Witz zeigt sich, dass es nur dann zu einem komplexen Verstehen und damit zum Lachen kommen kann, wenn der Leser alle Teilprozesse des Lesens erfolgreich durchlaufen hat und am Ende auf ein Mentales Modell zugreifen kann, in dem Textbasis und sprachlich-weltliches Wissen zusammengefasst sind. Dabei schließt sich der Bogen, wenn der Leser vom Schluss auf den Anfang zurückgreift und sieht, wie heikel solche großspurigen Wünsche sein können. Zu den Teilschritten gehört neben der Entnahme und Verarbeitung expliziter Informationen auf der Textbasis (Angler, drei Wünsche, Sarajewo) und anderer gradueller Tätigkeiten auf lokaler Ebene (z.B. das Erkennen der fiktiven Geschichte) eben auch die hierarchiehöhere Fähigkeit des Erkennens literarischer Mittel wie Witz oder Ironie, die im obigen Beispiel mit einem notwendigen Allgemeinwissen (Attentat und Erster Weltkrieg) verbunden werden müssen. Wie sich im Folgenden zeigen wird, bietet der DESI-Lesetest mit seinen hier skizzierten Kategorien und Stufen ein Instrumentarium, das es der Deutschlehrerschaft erlaubt, auf Individual- wie auf Klassenebene die jeweils erlangte Verstehenstiefe zu analysieren.

5.1 Testkonstrukt, Testaufgaben und Kompetenzniveaus

Der DESI-Lesetest fußt auf einem Prozessmodell, das sich neben kognitionspsychologischen ebenso aus linguistischen, literaturwissenschaftlichen und sprachpsychologischen Ansätzen speist, diese miteinander verbindet und sich somit von PISA 2000 unterscheidet (vgl. Willenberg 2004). Dabei wurde darauf geachtet, neben der Berücksichtigung der hierarchieniedrigen Teilkompetenzen des Lesens auch explizit zu machen, worauf sich hohe Lesekompetenzen stützen, nämlich auf die wissensbasierte Zusammenschau und Integration größerer Textteile, wohingegen PISA eher „Textverstehen im Nahbereich“ (Grzesik 2003, S.147), also eher auf lokaler Ebene und daher selten über Absätze hinweg getestet hat.

Damit die unten vorgestellten Testaufgaben in ihrer Schwierigkeit eingeschätzt werden können, beziehen sich die folgenden Hinweise auf die jeweiligen Umfänge

der zu verarbeitenden Textstellen. Aus Platzgründen können wir an späterer Stelle allerdings nur Ausschnitte von im Test verwendeten Texten präsentieren.

Ursprünglich wurden in DESI sechs aufeinander aufbauende Niveaus des Lesens angenommen, die auf verschiedenen Lesetätigkeiten beruhen und auf Basis der empirischen Ergebnisse schließlich zu den vier Lesekompetenzniveaus A-D zusammengezogen wurden. Im Folgenden sind die Schülerfähigkeiten auf diesen vier Niveaus kurz zusammengefasst. Eine Übersicht mit Aufgabenbeispielen gibt zusätzlich die Tabelle 5.1. Variieren auch die Textmengen, so bleiben doch die Teiltätigkeiten des Lesens (mit Abstrichen) die gleichen:

Niveau A: Einfache Sinnentnahme und Informationsverarbeitung (Identifizierende Lektüre)

- Die Fähigkeit, dem Text sinntragende Wörter zu entnehmen, d.h. basale Informationen über die Kerninhalte des verhandelten Themas, den Ort, die Zeit etc. ermitteln und verarbeiten zu können (vgl. auch PISA 2001). Textmenge: Einzelne oder mehrere Wörter.

Niveau B: Lokale Lektüre

- Die Fähigkeit, ad hoc *Inferenzen* zu bilden, d.h. an bestimmten Stellen lokale Schlussfolgerungen ziehen zu können, wenn zwischen zwei Sätzen oder innerhalb eines Satzes Begründungslücken oder vom Autor hinterlassene Leerstellen klaffen (vgl. auch Murray u.a. 1993). Textmenge: Ein oder zwei Sätze.
- Ferner: Die Fähigkeit zu *fokussieren*, d.h. den Lesefluss zu verlangsamen und die Aufmerksamkeit auf eine bestimmte, wichtige Stelle zu richten, um den Text auf diese Weise in thematische Konzepte abstrahierend strukturieren zu können (Willenberg 1999; Paefgen 2003). Textmenge: Eine Stelle aus einem oder mehreren Sätzen, in der lokale Kohärenz hergestellt wird.

Niveau C: Verknüpfungen

- Die Fähigkeit, externes (*allgemeines und textuelles*) Wissen heranziehen und für das Verstehen des Textes kritisch anzuwenden (vgl. auch Köster 2003). Außerhalb eines Tests gehört noch die Fähigkeit hierher, *Emotionssignale* innerhalb des Textes erkennen oder selbst lesebegleitend emotionale Muster aktivieren zu können. Textmenge: Die eingeführte Textthematik aufzunehmen kann i.A. nach einem oder mehreren Absätzen geschehen – außerhalb von Tests zuweilen auch bereits vor der Lektüre (durch Paratexte, PR-Maßnahmen etc.).
- Ferner: Die Fähigkeit, *Verknüpfungen* über zwei oder mehrere Absätze hinweg vorzunehmen, um auf diese Weise den „Faden“, d.h. das (zuweilen versteckte) Motiv des Texts erkennen zu können (vgl. Pütz 1998; Willenberg 1999; Beyer 2003). Textmenge: Über Absätze hinweg.

Niveau D: Mentales Modell

- Die Fähigkeit, ein *Mentales Modell* des gesamten Textes erstellen zu können (vgl. Kelter 2003). Mentale Modelle sind als konstruktive Zusammenfassung des Tex-

tes durch den Leser zu verstehen, die - bei literarischen Texten - aus den zentralen handelnden Figuren, dem Ort, der Zeit, dem jeweiligen Handlungskern sowie dem zentralen Motiv bestehen. Allerdings löst sich das Mentale Modell partiell von der Textbasis, da der Leser diese mit dem eigenen Vor- und Weltwissen elaborierend verbinden muss (vgl. Al-Diban 2002; Rickheit/Habel 1999), ehe es in seiner eigenen Modalität auf einer neuen (Meta)Ebene abgerufen werden kann.

Tabelle 5.1: Beschreibung der Kompetenzniveaus in Deutsch Lesen mit Aufgabenbeispielen. Schülerinnen und Schüler auf den höheren Kompetenzniveaus beherrschen jeweils auch die Inhalte der darunter liegenden Niveaus.

Kompetenz-niveau	Beschreibung	Beispiele
D (ab 651 Punkte)	Auswerten mentaler Modelle: Hauptfiguren und ihre Relationen, Ort, Zeit sowie ein zentrales Motiv zusammenfügen.	<i>Der Großvater sagt: „Durch die Brille sieht man auch den Schwindel größer.“ Was meint er damit?</i> Hier müssen die Relationen der beiden Personen und ihre zentralen Motive in Form eines Wortspiels erkannt werden.
C (635 bis 650 Punkte)	Verknüpfende Lektüre: Aktivieren von allgemeinem Wissen bzw. Textwissen; Verknüpfen zweier auseinander liegender Stellen, meistens um Motive oder Kausalitäten zu klären.	Textwissen anwenden (Literarischer Text): <i>„In der Geschichte gibt es einige komische bzw. lustige Stellen. Unterstreiche sie im Text!“</i> (Hier sollte das Textwissen angewandt werden, dass Komik meistens aus Erwartungsbrüchen oder aus Übertreibungen besteht.)
B (554 bis 634 Punkte)	Fokussierte Lektüre: Inferenzen bilden, besonders bei Lücken zwischen zwei Sätzen; genaues, fokussierendes Lesen an einer semantisch oder logisch schwierigen Stelle; Koordination von Zahlen in den Texten.	1. Frage zur Inferenzbildung (Literarischer Text): <i>„Warum meint die Großmutter, dass Bertka während ihrer Abwesenheit so alt geworden ist?“</i> (Die Schüler/innen müssen hier folgende Textstelle deuten.) <i>Sie besuchte ihre Klatschkumpankas und besah sie sich durch die [neue] Brille. „Gott, Bertka, was bist du alt geworden, die Zeit, wo ich nicht hier war!“</i> 2. Fokussieren, genau lesen (Sachtext): „Quastenflosser“: <i>„Wo hat man bisher die meisten lebenden Quastenflosser gefangen?“</i> a) Im Devon b) Bei den Komoren c) Im Meer d) Bei Afrika“ (Information im Text: Erster Fund bei Südafrika, mehrere weitere 3000 km entfernt bei den Komoren.)
A (322 bis 553 Punkte)	Identifizierende Lektüre: Erkennen der sinntragenden Einheiten in einem Satz oder einem Absatz.	Frage zur Identifikation sinntragender Wörter (Sachtext: „Quastenflosser“): <i>„Wovon handelt der Text?“</i> a) Von der Entwicklung der Fischerei im letzten Jahrhundert b) Vom Nachweis einer zoologischen Sensation (Unterstrichenes Zitat kommt wörtlich im Text vor.) c) Von der Tierwelt rund um Afrika d) Vom Aufbau und Aussehen eines Quastenflossers“

Da über das letzte Thema einige Unklarheit besteht (vgl. Bremerich-Vos/Grotjahn 2007) seien noch ein paar zusätzliche Bemerkungen angefügt. Seit der Erstveröffentlichung von Johnson-Laird zu diesem Thema (1983) ist auch im Deutschen eine Anzahl von Publikationen erschienen, die für das Leseverstehen hilfreich sind (Dutke 1998; Rinck 2000; Al-Diban 2002; Kelter 2003; Willenberg/Gailberger/Krelle 2007, Gailberger 2007a). Einigkeit besteht dabei, dass Mentale Modelle die Zusammenfassungen des Lesers sind, die nicht allein die explizite Textoberfläche repräsentieren, sondern darüber hinaus gehend ein subjektives Gesamtmodell bilden, das die Grundelemente und mehrere weiterführende (auch implizite) Erklärungen eines Textes integriert. Dabei wird die multiple Repräsentation des Speichermodells (Schiefele 1996) umso leichter und klarer, je weniger widersprüchliche Inferenzen vom Leser verarbeitet werden müssen (vgl. dazu Johnson-Laird 2001). Auf der anderen Seite aber verlangt es ein solches Modell auch, vorhandene Widersprüche und offene Möglichkeiten zu berücksichtigen, die dann für weitere Argumentationen unerlässlich sind (s.a. Radvansky/Copeland 2004). Im Deutschunterricht kann man diese Offenheit wesentlich intensiver erfassen, wenn man die Schüler/innen nicht zu einer klassischen Inhaltsangabe auffordert (dies sollte in diesem Sinne vielmehr vermieden werden), sondern ihnen z.B. die Möglichkeit gibt, die Entwicklung eines Textes narrativ zu beschreiben oder begründet darüber zu diskutieren. Damit berücksichtigt man einerseits die (mögliche) Dynamik Mentaler Modelle (vgl. Kelter 2003) und kommt so andererseits auch leichter an die subjektiven und zugleich logischen Verarbeitungen der Schüler heran, v.a. wenn man genau und argumentierend damit umgeht. Natürlich zeigen sich schon zu Lektürebeginn (je nach Vorwissen) erste Kernpunkte einer mentalen Modellbildung, aber erst am Ende und erst bei gelungenem Textverstehen bildet sich ein kohärentes Gerüst, das auf eine elaborierende Verarbeitung des situativen Leseaktes schließen lässt - jedes einigermaßen ausführliche Unterrichtstranskript beweist dies deutlich (vgl. hierzu Willenberg 2007).

Weitere Beispiele und Prozentanteile richtiger Lösungen der Aufgabenbeispiele

Niveau A, einfache Sinnentnahme: Erkennen der sinntragenden Einheiten in einem Satz oder einem Absatz. Sachtext: „Quastenflosser“

Frage: „Wovon handelt der Text? (74%)“

- a) Von der Entwicklung der Fischerei im letzten Jahrhundert
- b) Vom Nachweis einer zoologischen Sensation
- c) Von der Tierwelt rund um Afrika
- d) Vom Aufbau und Aussehen eines Quastenflossers“

Aufklärung: Die Formulierung von b) kommt wörtlich im Text vor, zwei weitere Distraktoren sind völlig falsch, Aussage d) trifft nur einen kleinen Teil des Textes.

Niveau B, Lokale Lektüre: Wie beschrieben, geht es hier um zwei Lesetätigkeiten; zum einen Inferenzen bilden, besonders bei Lücken zwischen zwei Sätzen, zum an-

deren um fokussierendes Lesen an einer semantisch oder logisch schwierigen Stelle; bzw. Koordination von Zahlen in Sachtexten.

Beispiel zu Inferenzen bilden: Literarischer Text „Großmutter“ (E. Strittmatter) „*Warum meint die Großmutter, dass Bertka während ihrer Abwesenheit so alt geworden ist?*“ (58%)

Die Schüler müssen hier nachfolgende Textstelle deuten, in der die nach nur kurzer Zeit zurückgekehrte, aber neu bebrillte Großmutter folgendes tut: „*Sie besuchte ihre Klatschkumpankas und besah sie sich durch die [neue] Brille. „Gott, Bertka, was bist du alt geworden, die Zeit, wo ich nicht hier war!“ [...]*“

Beispiel zum Fokussieren und zum Genauen Lesen: Sachtext „Quastenflosser“ „*Wo hat man bisher die meisten lebenden Quastenflosser gefangen?* (53%)

- a) Im Devon
- b) Bei den Komoren
- c) Im Meer
- d) Bei Afrika“

Es gibt mehrere Information im Text, die abgeglichen werden müssen: Erster Fund war im *Meer*, und zwar bei *Südafrika*, andere (und d.h. *mehrere*) aber 3000 km entfernt bei den Komoren.

Niveau C, Verknüpfende Lektüre: Auch hier gibt es wieder zwei Lesetätigkeiten, nämlich das Aktivieren von allgemeinem Wissen, bzw. von Textwissen und das Verknüpfen zweier auseinander liegender Stellen, meistens um Motive zu erkennen oder Kausalitäten zu klären.

Beispiel zum Anwenden von Textwissen (erneut „Großmutter“): „*In der Geschichte gibt es einige komische bzw. lustige Stellen. Unterstreiche sie im Text!*“ (36%)

Hier sollte das im Unterricht (laut Lehrpläne) behandelte Textwissen angewandt werden, dass Komik meistens aus Erwartungsbrüchen oder aus Übertreibungen entsteht.

Für ein **Beispiel zum Verknüpfen verschiedener Textstellen** sei unser Einleitungstext herangezogen, bei dem die Frage lauten könnte: *Beurteile, was der Sinn der Schlusspointe sein kann.* Für die Antwort müsste der Bezug zum Eingangswunsch hergestellt werden.

Niveau D, Bilden des Mentalen Modells: Das Zusammenfügen der Hauptfiguren in ihren Relationen, des Ortes, der Zeit, des Handlungskerns sowie eines zentralen Motivs zu einem Speichermodell. (Beispiel ist ein literarischer Text, wieder „Großmutter“): *Der Großvater sagt: „Durch die Brille sieht man auch den Schwindel größer.“ Was meint er damit?* (16%)

Hier muss auf der Basis des bisherigen Geschehens im Text das Wortspiel mit der Doppeldeutigkeit von Schwindel (Lüge und Unwohlsein), sowie die Relationen der beiden Personen und ihre zentralen Motive erkannt werden.

5.2 Ergebnisse

Leistungsverteilung am Ende der neunten Jahrgangsstufe

Abbildung 5.1 zeigt die Verteilungen der Schülerinnen und Schüler auf die Kompetenzniveaus in Deutsch Lesen am Ende der neunten Jahrgangsstufe.

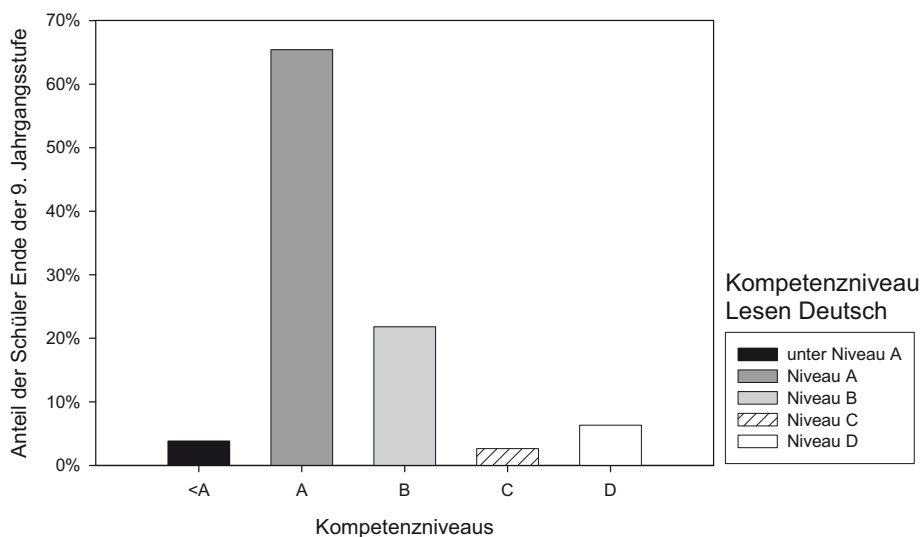


Abbildung 5.1: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Deutsch Lesen am Ende der neunten Jahrgangsstufe.

Ca. 4% der Schülerinnen und Schüler bleiben unter Niveau A und beherrschen damit die einfache Informationsentnahme nur zum Teil. Der Schwerpunkt der Leistungen liegt in der sicheren Beherrschung des Niveaus A. Insgesamt erfüllen am Ende der neunten Jahrgangsstufe 96% aller Schülerinnen und Schüler die oben erläuterten Anforderungen des Lesekompetenzniveaus A. 32% kommen darüber hinaus durch das Bilden von Inferenzen oder durch genauere Lektüre auf das Niveau B oder höher, während das Niveau C oder D noch 10% erreichen. Diese Schülerinnen und Schüler können bezogen auf die Testaufgaben externes Wissen anwenden und einer semantischen Spur innerhalb des Textes folgen. Das Mentale Modell (Niveau D) wird von 6% aller Schülerinnen und Schüler erfolgreich gebildet.

Leistungsveränderung im Laufe der neunten Jahrgangsstufe

Vergleicht man die gemessenen Unterschiede an den beiden Testzeitpunkten, so kann konstatiert werden, dass die Lesekompetenz der deutschen Schülerinnen und Schüler im Laufe der neunten Jahrgangsstufe nicht signifikant zunimmt. Abbildung 5.2 zeigt die Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Deutsch Lesen zu Beginn und am Ende der neunten Jahrgangsstufe. Wie ersichtlich, können sich allein die Gymnasiasten innerhalb der neunten Klasse verbessern.

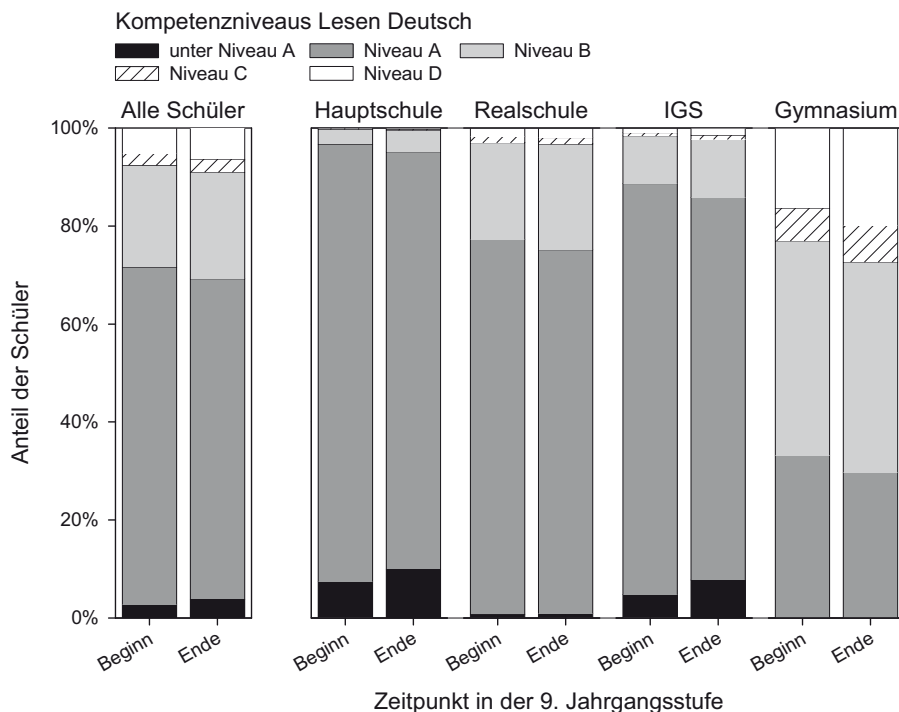


Abbildung 5.2: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Deutsch Lesen zu Beginn und am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Wenn man die Zuwächse in den Schularten vergleicht, sieht man, dass lediglich im Gymnasium ein erkennbarer Zuwachs und in der Realschule ein leichter zu erkennen ist. In der Hauptschule und der IGS nimmt nach einem Jahr sogar das Niveau unter Stufe A noch zu.

Geschlechterdifferenzierung und Zuwächse

Geschlechterdifferenzierung

In Tabelle 5.2 sind die Mittelwerte der Lesekompetenzen von Jungen und Mädchen sowie die Geschlechterdifferenzen aufgelistet. Wie ersichtlich, kamen die Mädchen mit dem DESI-Lesetest besser zurecht als die Jungen. Die Leistungen der Mädchen sind insgesamt signifikant besser als die der Jungen, dieser Unterschied ist außer in Hauptschulen und IGS auch in den einzelnen Bildungsgängen signifikant. Insgesamt spiegeln die Ergebnisse von DESI damit aber nicht allein die aktuelle Forschungslage zur besseren Lage der Mädchen beim Lesen wider. Denn negativ gewendet deutet dies augenscheinlich auf eine Jungenproblematik im Deutschunterricht, was schließen lässt, dass dieser bildungsgangübergreifend jungenspezifischer auszurichten sein wird als bisher geschehen (vgl. dazu Garbe 2007; Müller-Walde 2005; Eggert/Garbe 2003). Eine diesbezügliche Überarbeitung der Lehrpläne scheint hier dringend an-

geraten zu sein, die die neuropsychologischen und motivationspsychologischen Erkenntnisse mit den empirischen Daten der Lesesozialisationsforschung zu verbinden weiß.

Tabelle 5.2: Lesekompetenzen für Jungen und Mädchen sowie Geschlechtsdifferenzen am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

		Mädchen	Jungen	Differenz
Gesamt	M	510.9	488.8	22.2*
	SE	(3.20)	(3.41)	(3.83)
Hauptschule	M	426.0	417.7	8.3
	SE	(6.56)	(5.77)	(6.36)
Realschule	M	510.8	493.6	17.2*
	SE	(3.41)	(4.05)	(3.19)
IGS	M	460.8	450.6	10.1
	SE	(12.59)	(11.29)	(9.84)
Gymnasium	M	596.7	583.6	13.1*
	SE	(3.71)	(4.98)	(3.77)

Anmerkung: * Signifikant mit $p \leq .01$

Zuwächse der Leseleistungen nach Geschlechtern in den Bildungsgängen

Zusätzlich zu den in Tabelle 5.2 dargestellten Mittelwerten sind in Tabelle 5.3 die Zuwächse in Deutsch Lesen nach Bildungsgang und Geschlecht aufgelistet. Die Mädchen können (mit Einschränkung auf die IGS) ihr Lesen innerhalb der neunten Jahrgangsstufe signifikant verbessern, während dies den Jungen nicht gelingt – deren Lesekompetenz stagniert. Den höchsten signifikanten Zuwachs können wir bei den weiblichen Gymnasiasten beobachten (+11 Punkte), den höchsten Verlust bei den männlichen Hauptschülern (-7 Punkte) – wobei die deskriptiv negativen Zuwächse bei den Jungen nicht signifikant unter null liegen.

Tabelle 5.3: Veränderungen in Deutsch Lesen zwischen dem Beginn und Ende der neunten Jahrgangsstufe nach Geschlecht und Bildungsgang.

		Mädchen	Jungen	Differenz
Gesamt	M	7.5	-2.9	10.3*
	SE	(1.97)	(2.23)	(2.15)
Hauptschule	M	5.0	-7.2	12.2*
	SE	(4.01)	(4.66)	(4.52)
Realschule	M	8.7	-3.1	11.9*
	SE	(2.92)	(3.43)	(3.09)
IGS	M	-2.2	-0.6	-1.6
	SE	(6.17)	(6.58)	(7.32)
Gymnasium	M	10.8	1.7	9.1*
	SE	(3.94)	(4.57)	(3.61)

Anmerkung: * Signifikant mit $p \leq .01$

Differenzierung nach Erstsprache

Wie in Abbildung 5.3 veranschaulicht, lesen Schülerinnen und Schüler mit Erstsprache Deutsch besser als jene, deren Erstsprache gemischt („mehrsprachig“, vgl. Kap. 20) oder nicht deutsch ist. Dieses Ergebnis konnte erwartet werden. Betrachtet man die diesbezüglichen Ergebnisse allerdings genauer, so kristallisieren sich Ergebnisse heraus, die überraschen: So kann DESI zeigen, dass zweisprachige Gymnasiasten und auch solche mit einer nicht deutschen Erstsprache besser lesen als Gleichaltrige mit deutscher Erstsprache in der Real-, Gesamt- oder Hauptschule.

Ähnliche Beobachtungen, bei denen Schülerinnen und Schüler mit nicht deutschem oder gemischtem Sprachhintergrund höhere Sprachkompetenzen im Deutschen nachweisen als ihre gleichaltrigen deutschen Kompagnons, können auch in den Modulen Wortschatz und Argumentation gemacht werden.

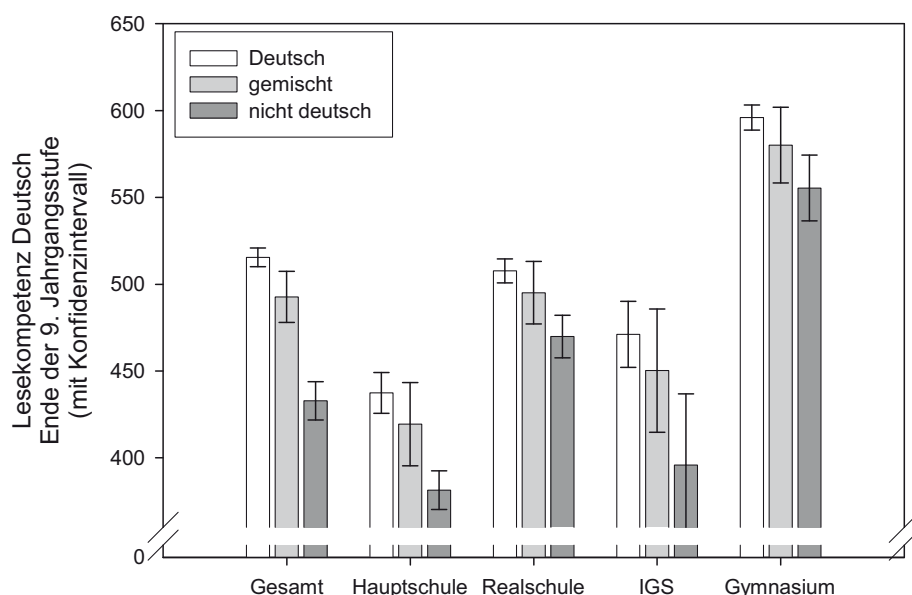


Abbildung 5.3: Mittelwerte in Deutsch Lesen nach Erstsprache und Bildungsgang.

Im Gymnasium sind die Leseleistungen aller drei sprachlichen Gruppen besser als in den anderen Bildungsgängen. Ansonsten übertreffen die deutschsprachigen Schülerinnen und Schüler die anderen signifikant. Verbesserungsbedürftig sind die Leistungen der fremdsprachigen Jugendlichen insbesondere in der Hauptschule und in der IGS.

Besondere Aufmerksamkeit sollte daher in den kommenden Jahren der deutsch- bzw. lesedidaktischen Forschung solcher Risikogruppen wie den Hauptschülerinnen und Hauptschülern mit nicht deutscher Erstsprache zukommen (vgl. dazu nur Rosebrock/Nix 2008). Sowohl ihre Daten bezüglich der zum zweiten Testzeitpunkt erreichten Lesekompetenz als auch der bei ihnen gemessene si-

gnifikante Leistungsrückgang innerhalb eines Jahres von -12.2 Punkten zeigt dies überdeutlich.

5.3 Perspektiven und Konsequenzen für Deutschdidaktik und Deutschunterricht

Positiven Werten oder Zuwächsen in der Lesekompetenz einer Gruppe stehen häufig niedrige Werte oder gar Verluste einer anderen gegenüber (z.B. Mädchen vs. Jungen, Gymnasium vs. Integrierte Gesamtschule, mit deutscher Erstsprache vs. mit nicht deutscher Erstsprache). Es scheint daher angezeigt, dass schulart-, geschlechts- und sprachspezifischer als bisher unterrichtet werden sollte, um die von DESI aufgezeigten Probleme zu bearbeiten und die Lücken zu schließen. Dies reicht von Fragen bezüglich des alltagsüblichen Deutschunterrichts bis hin zur Verfeinerung der Curricula und der Bildungsstandards.

Zur Schularbeit: Neben den soliden wenn auch verbesserungswürdigen Ergebnissen der Gymnasiasten und der Realschüler fallen v.a. die schlechteren Ergebnisse in den Hauptschulen und den Integrierten Gesamtschulen auf. Hier kann vor dem Hintergrund der eingangs erläuterten und mit den Testergebnissen kompatiblen DESI-Leseprozess Theorie ein Literatur- und Lesecurriculum entwickelt werden, in dem alle nötigen Teilanforderungen des Lesens von der Informationsentnahme bis zum Mentalen Modell des Schülers Berücksichtigung finden können. Besonders sei an dieser Stelle die Rolle des (Vor)Wissens herausgestellt, dem wie oben erwähnt auch in anderen Modulen von DESI eine tragende Rolle in der Sprachkompetenz zugeschrieben werden konnte (vgl. hierzu nur Gailberger/Krelle/Triebel 2007).

Zu den Jungen: Wir haben es mit einer Jungenproblematik von besonderer Art im lesespezifischen Deutschunterricht zu tun, die über den uns bekannten ‚Leseknick‘ hinausgeht. Hier scheinen Genre- und Themenaufwertungen für die Jungen dringend angezeigt, etwa wenn es um die weitere Etablierung von Sachtexten (Geschichte, Naturwissenschaft, Technik, Sport), um Helden- oder Bandengeschichten geht, die es den Jungen jeweils erlauben, sich mit den Protagonisten, d.h. mit ihren Stärken, ihren Schwächen und ihren Entwicklungen zu identifizieren. So könnte es ihnen unter Berücksichtigung ihrer Lese-Interessen und je eigenen Zugängen zur Welt möglich werden, sich wirklich gründlich mit Texten auseinander zu setzen (vgl. dazu nur Gailberger 2007b).

Zur deutschen Sprache: Die Ergebnisse des DESI-Tests des Lesens zeigen die Dringlichkeit auf, mit der konkrete Leseförderungsmaßnahmen für in Deutschland lebenden Schüler nicht deutscher Sprachlichkeit entwickelt werden müssen. Dabei sind es v.a. die Haupt- und Gesamtschüler mit nicht deutscher Erstsprache, bei denen häufig Basisarbeit geleistet werden muss. Hier können Methoden und Forschungsarbeiten aus der angelsächsischen Leseforschung etwa zur Bedeutung und Förderung der *Reading Fluency* (vgl. Rosebrock/Nix 2006), aus den Fremdsprachendidaktiken zum simultanen Einsatz von Hörmedien beim Lesen (vgl. Hinz 1992; sowie Gailberger 2007c für die Deutschdidaktik) wie aus dem Bereich Deutsch als Zweit-

bzw. Fremdsprache (vgl. Helbig u.a. 2001) als erste Anleihen dienen, die auf den Deutschunterricht übertragen werden könnten. Hier erscheinen auch die verstärkten Arbeiten mit Wörterlisten, mit Bildwörterbüchern wie auch mit verbindlichen Vokabelheften zum Ausbau und Sichern des Basiswortschatzes als notwendig, damit diese Schülerinnen und Schüler aktiv am gesellschaftlichen und ökonomischen Leben in der Deutschland teilhaben können.

Literatur

- Al-Diban, S. (2002): Diagnose mentaler Modelle. **Hamburg: Dr. Kovac.**
- Beyer, R. (2003): Verstehen von Diskursen. In: Rickheit, G./ Herrmann, T./Deutsch, W. (Hrsg.): Psycholinguistik. Ein internationales Handbuch (HSK), Berlin und New York, S.532-543.
- Bremerich-Vos, A./Grotjahn, R. (2007): Lesekompetenz und Sprachbewusstheit. In: Beck, B./ Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 158-177.
- Dutke, S. (1998): Zur Konstruktion von Sachverhaltsrepräsentationen beim Verstehen von Texten: 15 Jahre nach Johnson-Lairds ‚Mental Models‘. In: Zeitschrift für Experimentelle Psychologie 45, S.42-59.
- Eggert, H./Garbe, Ch. (2003): Literarische Sozialisation, Stuttgart: Metzler
- Gailberger, St. (2007a): Die Mentalen Modelle der Lehrer elaborieren. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht, Hohengehren: Schneider.
- Gailberger, St. (2007b): Lesen und Lesebewusstheit im Deutschunterricht. Versuch einer multi-dimensional basierten Leseprozessstheorie zur Förderung des Lesens im Deutschunterricht. In: Gailberger, St./Krelle, M. (Hrsg.): Wissen und Kompetenz. Entwicklungslinien und Kontinuitäten in Deutschdidaktik und Deutschunterricht. Baltmannsweiler: Schneider, S.149-173.
- Gailberger, St. (2007c): Leseförderung durch Hörbücher – Eine verbal-auditive Leseförderungstheorie für den Deutschunterricht. In: Lecke, B. (Hrsg.): Medien im Deutschunterricht, Frankfurt/ M u.a.: Peter Lang, (In Druck).
- Gailberger, St./Krelle, M./ Triebel, W. (2007): Wissen und Kompetenz im Literaturunterricht am Beispiel von Nachts schlafen die Ratten doch. In: Gailberger, St./Krelle, M. (Hrsg.): Wissen und Kompetenz. Entwicklungslinien und Kontinuitäten in Deutschdidaktik und Deutschunterricht. Baltmannsweiler: Schneider, S. 97-119.
- Garbe, C. (2007): Lesen – Sozialisation – Geschlecht. Geschlechterdifferenzierende Leseforschung und –förderung. In: Bertschi-Kaufmann, A. (Hrsg.): Lesekompetenz – Leseleistung – Leseförderung: Grundlagen, Modelle und Materialien. Seelze: Kallmeyer, S.66-82.
- Grzesik, J. (2003): Was testet der PISA-Test des Lesens? In: Abraham, U./Bremerich-Vos, A./ Frederking, V./Wieler, P. (Hrsg.): Deutschdidaktik und Deutschunterricht nach PISA. Freiburg: Fillibach. S.135-164.
- Helbig, G./Götze, L./Henrici, G./Krumm, H.-J. (Hrsg.) (2001): Deutsch als Fremdsprache. Ein internationales Handbuch (HSK). Berlin, New York: Walter de Gruyter.
- Hinz, K. (1992): Hörspiele in der leistungsschwachen Englischklasse. In: *Praxis des neusprachlichen Unterrichts* 4, S. 372-378.
- Johnson-Laird, Ph. N. (1983): Mental models: Towards a cognitive science of language, inference, and consciousness. Cambridge: CUP.
- Johnson-Laird, P. N. (2001): Mental models and deduction. In: *Trends in Cognitive Sciences* 5 (10), S.434-442.

- Kelter, St. (2003): Mentale Modelle. In: Rickheit, G./Herrmann, T./Deutsch, W. (Hrsg.): Psycholinguistik. Ein internationales Handbuch (HSK). Berlin, New York: Walter de Gruyter, S. 505-517.
- Köster, J. (2003): Die Bedeutung des Vorwissens in der Lesekompetenz. In: Abraham, U./Bremerich-Vos, A./Frederking, V./Wieler von Fillibach, P. (Hrsg.): Deutschdidaktik und Deutschunterricht nach PISA. Freiburg i. Brsg.: Fillibach, S. 90-105.
- Müller-Walde, K. (2005): Warum Jungen nicht mehr lesen und wie wir das ändern können. Frankfurt a/M: Campus.
- Murray, J.D./Klin, C.M./Myers, J. (1993): Forward inferences in narrative textes. In: *Journal of Memory and Language* 32, S. 464-473.
- Paefgen, E. (2003): Textnahes Lesen und Rezeptionsdidaktik. In: Kämper van den Boogaart, M. (Hrsg.): Deutsch Didaktik - Leitfaden für die Sekundarstufe I und II. Berlin: Cornelsen Verlag Scriptor, S. 191-209.
- PISA-Konsortium Deutschland (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske & Budrich.
- PISA-Konsortium Deutschland (Hrsg.) (2004): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse der zweiten internationalen Studie. Münster: Waxmann.
- Pütz, D. (1998): Lehren, Lernen, Organisieren. Baltmannsweiler: Schneider.
- Radvansky, G.A./Copeland, D.E (2004): Reasoning, Integration, Inference and Text Comprehension. In: *Canadian Journal of Experimental Psychology* 58 (2), S. 133-142.
- Rickheit, G./Habel, Ch. (Hrsg.) (1999): Mental Models in Discourse Processing and Reasoning. Amsterdam: Elsevier.
- Rinck, M. (2000): Situationsmodelle und das Verstehen von Erzähltexten: Befunde und Probleme. In: *Psychologische Rundschau* 51 (3), S. 115-122.
- Rosebrock, C./Nix, D. (2006): Forschungsüberblick: Leseflüssigkeit (Fluency) in der amerikanischen Leseforschung und -didaktik. In: *Didaktik Deutsch* 20, S. 90-111.
- Rosebrock, C./Nix, D. (2008): Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung. Baltmannsweiler: Schneider.
- Schiefele, U. (1996): Motivation und Lernen mit Texten. Göttingen: Hogrefe.
- Willenberg, H. (1999): Lesen und Lernen – Eine Einführung in die Neuropsychologie des Textverstehens. Heidelberg, Berlin: Spektrum.
- Willenberg, H. (2004): Lesestrategien. In: *Praxis Deutsch* 187, S. 6-15.
- Willenberg, H. (Hrsg.) (2007): Kompetenzhandbuch für den Deutschunterricht. Auf der empirischen Basis des DESI-Projektes. Baltmannsweiler: Schneider.
- Willenberg, H./Gailberger, S./Krelle, M. (2007): Argumentation. In: Beck, B. & Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 118-129.

Heiner Willenberg

6 Wortschatz Deutsch

Das Testkonzept beruht auf ausgewählten linguistischen Theorien, die folgende Aspekte umfassen:

- Der Wortschatz von Menschen ist nach Sachfeldern wie auch nach Wortfeldern gespeichert.
- Die Schwierigkeiten lassen sich durch eine Skala kennzeichnen, deren Pole von großer Häufigkeit des Vorkommens bis zu seltenem Gebrauch reichen, dabei spielen Konkretion und Abstraktion eine Rolle.
- Die Bedeutungen von Wörtern bestimmen sich auch durch die Nuancen des Kontextes, in denen sie verwandt werden.

Auf Basis der Testauswertung ergaben sich drei Kompetenzstufen für den Wortschatz, die in Beispielen vorgestellt werden. Dieses Modul ist das einzige in der Studie, in dem die Jungen gleich gute Leistungen wie die Mädchen erreichten. Als bedeutsam erweisen sich die Resultate in den Schularten, wobei sich eine Dominanz des Gymnasiums ergibt, dort erzielen sogar zweisprachige Schüler/innen bessere Ergebnisse als alle Jugendlichen in den anderen Bildungsgängen. Weiterhin bemerkenswert ist der generelle Vorsprung der deutschsprachigen Schüler gegenüber denjenigen mit einer nicht deutschen Erstsprache, der mindestens eine Kompetenzstufe ausmacht. Eine angemessene Behandlung des Wortschatzes für alle Schüler, sowohl für die deutschen Muttersprachler wie für die Migranten, hat der Unterricht mitsamt der Didaktik bisher vernachlässigt.

6.1 Testkonstrukt und Testaufgaben

Es gibt überraschenderweise kaum Untersuchungen über die Differenzierung und den Reichtum des Wortschatzes bei Jugendlichen. Die Publikationen zur Jugendsprache untersuchen lediglich die besonderen Eigenarten in der Kommunikation dieser Altersgruppe und ihrer speziellen Ausdrücke (vgl. Schlobinski 1998) oder sie behandeln partielle Einzelthemen (vgl. Augst 1974). Aus diesem Grund haben wir die didaktisch relevanten Publikationen in der Wortschatzforschung durchsucht, um Hinweise für Testkonzepte zu gewinnen (vgl. Willenberg 2007). Als hilfreich und verträglich mit den Lehrplänen erwiesen sich folgende Ansätze:

a) Die Wortfeldforschung: Jedes Lexem fügt sich in ein Netz von verwandten und nuancierenden Einträgen ein. Unmittelbar dazu gehört die Theorie vom Sachfeld, dessen Einträge sich gleichsam aus der Praxis der Sprecher ableiten, die in vielen realen Situationen eine Fülle von genauen Bezeichnungen beherrschen müssen. Dies gilt sowohl bei alltäglichen Dingen wie z.B. einer Wohnungseinrichtung als auch bei speziellen Fachgesprächen wie über den Wald oder über einen Motor (vgl. Schmidt 1973). Es ist inzwischen auch psychologisch nachgewiesen worden, dass sich das

mentale Lexikon von Menschen nach Sachgebieten ordnet, allerdings individuell je unterschiedlich (Damasio/Damasio 1992).

b) Die Satzsemantik: Wörter und Ausdrücke erhalten ihren nuancierten Sinn in syntaktischen Kontexten, so dass man erst dort Ironie, Wertungen oder Metaphorik erkennen kann (Polenz 1988).

c) Untersuchungen zu den Ebenen des Wortschatzes und seiner Häufigkeit (vgl. Miller 1993): Diese Sichtweise ist u.a. mit dem Merkmalskonzept verbunden, das besagt, Wörter sind in ihrer linguistischen Analyse durch eine gut erkennbare Anzahl von definierenden Aspekten zu beschreiben. Man kann eine Skala herstellen, die von häufig gebrauchten Konkreta bis zu selteneren Abstrakta und Fachwörtern reicht. Am schwierigen Ende der Skala werden die Einträge meist in speziellen Zusammenhängen gebraucht und sie enthalten teilweise Merkmale, die sich erst durch Definitionen und Beispiele erschließen lassen.

d) Einen stärker der Sprachproduktion verpflichteten Ansatz, wie ihn z.B. Aitchison (1997) darstellt, konnten wir im Rahmen dieses Tests nicht entfalten. Der DESI-Wortschatztest umfasst Testteile, die auf die genannten Theorien Bezug nehmen und Wortfelder aktiv erfragen, Merkmale oder Definitionen verlangen, Einträge in Satzkontexte erfordern oder anhand von Sachfeldern die Bezeichnungen von Gegenständen abfragen.

6.2 Niveaus der Wortschatzkompetenz

Es ergaben sich im DESI-Test drei Niveaus, deren Bezeichnungen am besten durch die Termini der Ebenentheorie bestimmt werden, die auch indirekt etwas über die Häufigkeiten der jeweiligen Wörter aussagt, nämlich:

- Häufig vorkommende Einträge im Grundwortschatz (definiert nach den ersten 2000 Wörtern im Langenscheidt Grundwortschatz (1990).
- Häufigere Konkreta oder Abstrakta, die nicht zum Grundwortschatz gehören.
- Seltener Fach- oder Fremdwörter sowie übertragene Redensarten, die aber zum Wissensgebiet von Neuntklässlern gehören können.

In Tabelle 6.1 sind die Kompetenzniveaus, deren Beschreibungen und Beispielimens dargestellt.

Tabelle 6.1: Beschreibung der Kompetenzniveaus in Deutsch Wortschatz mit Aufgabenbeispielen.

Kompetenz-niveau	Beschreibung	Beispiele
C (ab 595 Punkten)	Seltenere und spezielle Fach- oder Fremdwörter sowie Redewendungen, die zum Wissensgebiet von Neuntklässlern gehören können	<p>Gib dem Bild einen Titel und bezeichne die nummerierten Gegenstände auf dem Bild so genau wie möglich! Umschreibungen sind nicht erlaubt.</p> <p>TITEL: _____</p> <p>Die geforderte Überschrift ist z.B. „Flohmarkt“, die erfragten Gegenstände sind im Falle des Bildes „Flohmarkt“ zu benennen mit z.B: Abakus.</p> <p>Modedesigner Hedi Slimane: „Ich muss sehen, wie es aussieht, wenn jemand meine Kleider trägt. Dabei geht es immer mehr um die Beobachtung als um die Interpretation – eine eher emotionale als ____?__?__?__?__ Arbeit.</p> <p>a) intellektuelle; b) intelligente; c) vernünftige; d) denkerische</p>
B (550 bis 594 Punkte)	Häufigere Konkreta oder Abstrakta, die nicht zum Grundwortschatz gehören, aber im Sprachgebrauch von Neuntklässlern vorkommen.	Das Wort Standuhr wird aktiv als Bezeichnung in die Leerstelle eines Bildes eingefügt, das einen Flohmarkt zeigt.
A (473 bis 549 Punkte)	Häufig vorkommende Einträge im Grundwortschatz: Die ersten 2000 Einträgen im Grundwortschatz von Langenscheidt (1991).	<p>Finde genauere Ausdrücke an den Stellen, an denen Wörter durchgestrichen sind: Sagen</p> <p>Mirjam sagte zu Frank: „Mit dir gehe ich nicht mehr zum Konzert. Denn du hast mich vor meiner besten Freundin schlecht gemacht.“</p> <p>a) Frank sagte: „Das stimmt überhaupt nicht. Sie ist über dich hergezogen.“</p> <p>b) Mirjam sagte daraufhin: „Meine beste Freundin würde das niemals tun!“</p>

Weitere Beispiele und Prozentanteil richtiger Lösungen der Beispielaufgaben

Kompetenzniveau A:

Häufig vorkommende Einträge im Grundwortschatz

- Etwas auf der Zunge zergehen lassen: 89%. In einem Einfügungstest: Wohlbe- finden entsteht, wenn alle Sinne im Gleichklang angeregt werden. So lassen sich zum Beispiel zwei New Yorker Mädchen – schreiend vor Freude und mit weit aufgerissenen Augen den ersten richtigen Wintertag auf der Zunge ____?_?_?_?_? _____. Die Distraktoren waren: ausbreiten /verlaufen / zergehen /schmelzen.
- Sich kühl anfühlen: 82%. Freie Einsetzung zum Thema „Kälte“: „Er nahm ihre Hand, sie fühlte sich kühl an.“
- Meinen: 77%. In einem Kontext muss „sagen“ ersetzt werden: Die Schüler mussten „meinen“ selbständig finden.
- Ofen: 71%. Das Wort Ofen wird aktiv als Bezeichnung in die Leerstelle eines Bil- des eingefügt, das einen Flohmarkt zeigt.

Kompetenzniveau B:

Konkrete / Abstrakta außerhalb des Grundwortschatzes

- Standuhr: 51%. Das Wort Standuhr wird aktiv als Bezeichnung in die Leerstelle eines Bildes eingefügt, das einen Flohmarkt zeigt.
- chartern: 58% / defensiv: 67%. Finden der richtigen deutschen Bezeichnung; bei- de in Mehrfachwahlaufgabe.

Kompetenzniveau C:

Fachwörter / Fremdwörter und übertragene Redeweisen

- Matrosenanzug: 39%. Einsetzung im Flohmarktbild
- Abakus: 13%. Einsetzung im Flohmarktbild
- Intellektuelle Arbeit: 37%. Einfügungstext: Modedesigner Hedi Slimane: „Ich muss sehen, wie es aussieht, wenn jemand meine Kleider trägt. Dabei geht es im- mer mehr um die Beobachtung als um die Interpretation – eine eher emotionale als ____?_?_?_?_? ____ Arbeit.“ Die Distraktoren waren: intellektuelle / intelligente / vernünftige / denkerische
- Ein trojanisches Pferd benutzen (einsetzen): 36%. Der Geschäftsführer war ver- zweifelt: „Wir müssen wohl oder übel ____?_?_?_?_? ____ , um das Erfolgsgeheim- nis der Konkurrenz zu lüften.“ Die Distraktoren waren: den Rubikon überschrei- ten / Herkulesarbeit leisten / ein trojanisches Pferd einsetzen / Sisyphusarbeit verrichten.

6.3 Ergebnisse

Verteilungen der Schülerleistungen auf die Kompetenzniveaus am Ende der neunten Jahrgangsstufe

Der Entwicklungsstand des Wortschatzes präsentiert sich am Ende der neunten Jahrgangsstufe im Sinne der Leistungsverteilung aller Schülerinnen und Schüler auf die Kompetenzniveaus wie in Abbildung 6.1 dargestellt.

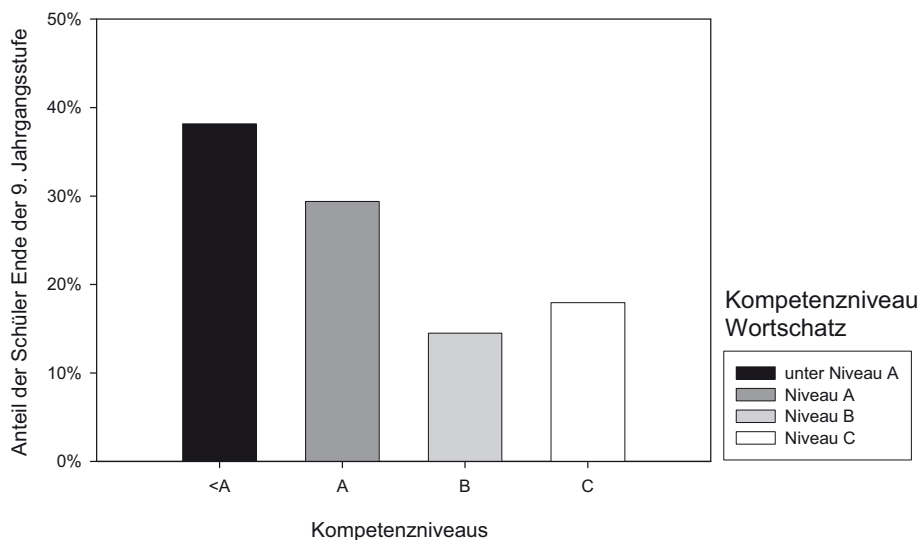


Abbildung 6.1: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus im Wortschatz am Ende der neunten Jahrgangsstufe.

Abbildung 6.2 zeigt zusätzlich, wie sich die Leistungen in den vier Bildungsgängen verteilen. In der Hauptschule beherrschen mehr als zwei Drittel den Grundwortschatz lediglich teilweise (unter Niveau A). In der Realschule befinden sich ca. 70% auf den Niveaus A-C. Auch die Abbildung der IGS zeigt mehr als die Hälfte der Schülerinnen und Schüler mit nur teilweise beherrschtem Grundwortschatz. Im Gymnasium erreichen die meisten Schüler das Niveau C.

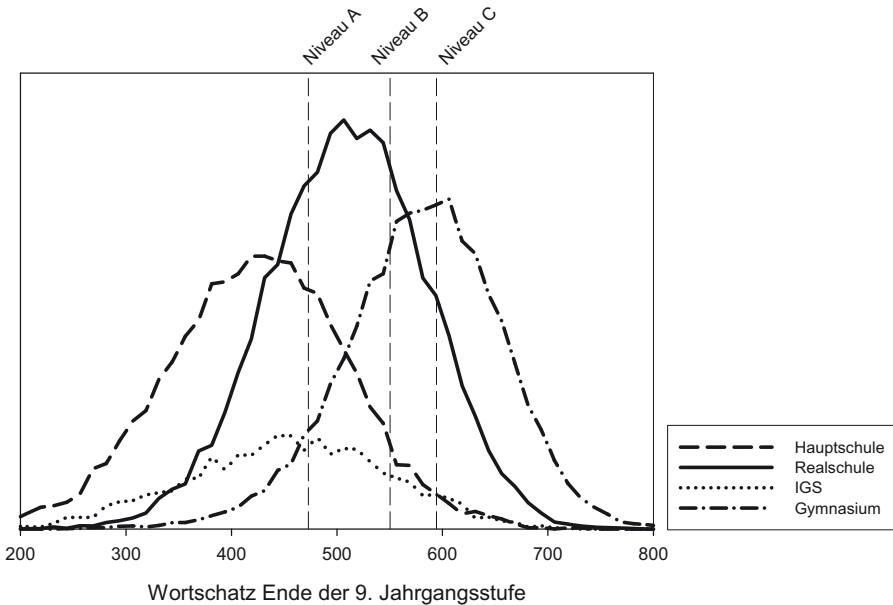


Abbildung 6.2: Die Ergebnisse des Wortschatztests in den Bildungsgängen am Ende der neunten Jahrgangsstufe.

Leistungsverteilung nach Sprache, Bildungsgang und Geschlecht

Aufschlussreich sind die Ergebnisse, wenn man sie nach der Erstsprache der Schülerinnen und Schüler aufschlüsselt. Zur Charakterisierung des sprachlichen Hintergrundes wird folgende Dreiteilung verwendet (vgl. Kapitel 20): Deutsch, Mehrsprachigkeit (zu Hause wird neben Deutsch mindestens eine weitere Sprache benutzt) und nicht Deutsch (eine nicht deutsche Sprache als Verständigungsmittel bzw. als angegebene Erstsprache).

Tabelle 6.2: Mittelwerte im Wortschatztest nach Sprache und Bildungsgang.

Sprache	Gesamt	Hauptschule	Realschule	IGS	Gymnasium
Deutsch	519.8	448.3	516.9	477.7	588.8
Mehrsprachig	478.6	396.7	495.8	444.0	556.2
Nicht Deutsch	407.7	358.5	455.7	364.9	508.0

Tabelle 6.2 zeigt die Leistungen der drei Sprachgruppen insgesamt und in den vier Bildungsgängen, wobei 500 den Mittelwert der gesamten Stichprobe darstellt. Man sieht, dass die deutschsprachigen Schüler und Schülerinnen in allen vier Bildungsgängen deutlich bessere Leistungen erzielen. Wenn man die Gesamtmittelwerte zwischen deutschsprachigen Jugendlichen und denen mit einer anderen Hauptsprache ansieht, dann beträgt die Differenz 112 Punkte. Auch in den einzel-

nen Schularten liegen die Differenzen zwischen 61 und 113 Punkten, darin liegen Differenzen, die den Unterschied mindestens eines Kompetenzniveaus ausmachen.

Es ist aber auch nicht zu übersehen, dass im Gymnasium die mehrsprachigen Jugendlichen (556) durchschnittlich besser sind als die „einheimischen Sprecher“ in den anderen drei Bildungsgängen. Das heißt, es findet sich hier eine Gruppe, die sich im Wortschatz so gut entwickelt hat, dass sie mit den Muttersprachlern in anderen Schulen mithalten kann. In die gleiche Richtung geht die Beobachtung, dass diejenigen Gymnasiasten, die eine andere Erstsprache als Deutsch angeben (508), immerhin die Sprachgruppierungen aller anderen Schularten übertreffen mit Ausnahme der deutschsprachigen Realschüler (517) - also auch hier ein Bildungserfolg in der Sprache unseres Landes.

Tabelle 6.3: Mittlere Leistungen der Jungen und Mädchen aufgeschlüsselt nach Sprache, Geschlecht und Bildungsgang.

Sprache	Geschlecht	Gesamt	HS	RS	IGS	Gymnasium
Deutsch	weiblich	524.8	450.3	516.7	487.6	587.7
	männlich	514.5	446.6	517.2	469.0	590.3
Mehrsprachig	weiblich	477.7	398.3	493.5	438.7	553.9
	männlich	479.9	394.6	498.4	451.6	559.5
Nicht Deutsch	weiblich	404.4	356.5	450.8	349.0	507.4
	männlich	410.8	360.5	460.2	376.1	508.5

Tabelle 6.3 fügt den Unterschieden in Sprachen und Schularten noch die Leistungen der Geschlechter hinzu, wobei sich zeigt, dass die Jungen in vielen Spalten minimal bessere Ergebnisse als die Mädchen vorweisen, die aber nicht signifikant werden. Auch der stärkste Vorsprung der Jungen mit nicht deutscher Erstsprache gegenüber den Mädchen in der IGS ist noch nicht signifikant.

Man kann in den Befunden auch einen Beitrag zur aktuellen Diskussion sehen, dass Jungen in sachlichen Zusammenhängen durchaus interessiert oder leistungsfähig sind. Wahrscheinlich müssen sie das Gefühl bekommen, Kompetenzen für das Leben oder wenigstens für eine Praxis zu erwerben. Die neuere Forschung zur Lernpsychologie Piagets bestätigt diese These durch Experimente: Kinder beherrschen kognitive Operationen leichter und früher als die bisherige Forschung es angenommen hatte, wenn ihr emotionaler Zugang zu den Themen positiv ist. Und dann können sie durchaus kognitive Grenzen überspringen (Houdé 2004).

6.4 Didaktische Interpretation der Ergebnisse und einige Folgerungen

Unklar ist die Beschäftigung der Lehrerschaft mit diesem Thema. Im Lehrerfragebogen erscheint in der Skala „Unterrichtsmittel Wortschatz“ ein relativ hoher Mittelwert, der über alle Bildungsgänge berechnet bei 3.1 Punkten liegt (auf einer Skala von 1-5).

In der Didaktik findet sich allerdings nur ein schwacher Strom von Publikationen zu diesem Thema, der sich außerdem auf wenige Aspekte bezieht. Pro Jahr erscheinen im Durchschnitt drei bis vier Artikel zur Wortschatzarbeit in allen drei Schulstufen. Am ehesten beziehen sich die Autoren auf die Grundschule und schlagen für die Kinder den Umgang mit Wörterbüchern und mit Wortfeldern vor, exemplarisch sei Peter Kühns Grundschriftwörterbuch genannt (Kühn 2005). Für die anderen Schulstufen erscheinen lediglich Solitäre, und nur wenige didaktische Übersichten beschäftigen sich damit (Löschmann 1993; Ulrich 2001).

So liegt die Vermutung nahe, dass die Lehrerschaft zwar immer wieder Wortschatzprobleme thematisiert, sie aber kaum in systematischer Weise behandelt und dabei bisweilen übersieht, wie wesentlich ein stabiler und differenzierter Wortschatz für alle Leistungen im Deutschunterricht ist. Auffällig ist jedenfalls die hohe Zahl von 38% der Schülerschaft, die unter dem Niveau 1 liegt, bei weitem die höchste Zahl in allen Deutschmodulen.

DESI hat dem Wortschatz im Orchester des Deutschunterrichts eine tragende Rolle zugestanden, was in deutschen Projekten selten ist und vielleicht auch einen der Gründe für die Vernachlässigung dieses Themas darstellt. In den Niederlanden fanden Schoonen/Hulstijn/Bossert bereits 1998, dass eine gute Kenntnis des Vokabulars der Erstsprache als einer der beiden Hauptprädiktoren für die Lesefähigkeit gelten kann. Sicherlich gibt es beide Wirkungsrichtungen: Intensives Lesen erweitert den Wortschatz und ein differenziertes, leicht zugängliches inneres Lexikon unterstützt die Lektüre (vgl. Anderson/Freebody 1985). Lesen wird allenthalben gefördert, sein Kompagnon, die gute Wortschatzbasis, bleibt leider meist im Halbdunkeln.

Einige Vorschläge für die Förderung dieses sprachlichen Handwerkszeuges sind in letzter Zeit gemacht worden, z.T. bereits in Kenntnis der DESI-Resultate:

- Peter Kühn führt Beispiele text- und kontextgestützter Differenzierung vor, die den Wortschatz im mentalen Lexikon verankern (2007).
- Daniel Nix plädiert für Vorleseübungen in Zweiergruppen und stützt sich dabei auf die phonologische Spur (2007).
- Heiner Willenberg ordnet diverse existierende Wortschatzübungen den drei Schwierigkeitsstufen dieses Projekts zu (2007).
- Winfried Ulrich propagiert, die Schülerschaft für die Lexik zu sensibilisieren, die mit einer Sprachreflexion einhergehen sollte (2007). Die relativ hohe Korrelation der beiden entsprechenden DESI-Module (vgl. Kapitel 18) könnte dafür eine argumentative Hilfe bieten.

Ein weiterer Vorschlag, den Wortschatz der Schüler zu vergrößern, ist, mit ihnen auch in Deutsch ein Vokabelheft zu führen, das als spannende individuelle Bereicherung betrachtet werden kann: Wörter, die jemand nicht versteht, schreibt er auf, sucht eine eigene Definition und fügt dann einen Beispielssatz hinzu. Und diese Einträge werden dann in einem Wettbewerb ab und zu wiederholt und damit trainiert, so dass sie als Vademecum deklariert werden können. Wenn den Schülern die Wichtigkeit des Wortschatzes klar ist, dann reagieren nach Erfahrungen auch Jungen positiv auf diese Methode. Eine andere Variante besteht darin, Wissensnetze zu einzelnen Themen

zu sammeln und sie auch – wenn es geht – optisch zu dokumentierten bzw. sie mit neueren grafischen Methoden auf einer Seite zu versammeln (s. dazu auch Rudolph 2007).

Literatur

- Aitchison, J (1997): Wörter im Kopf. Eine Einführung in das mentale Lexikon. Tübingen: Max Niemayer.
- Anderson, P.A./Freebody, P. (1985): Vocabulary knowledge. In: Singer, H./Ruddell, R.B. (Hrsg.): Theoretical models and processes of Reading. Newark: IRA, S. 343-371.
- Augst, G. (Hrsg.) (1974): Spracherwerb von 6 bis 16. Düsseldorf: Schwann.
- Damasio, A. R./Damasio, H. (1992): Sprache und Gehirn. In: Gehirn und Bewußtsein. Heidelberg: Spektrum, S. 395-398.
- Detering, K. (Hrsg.) (2000): Wortschatz und Wortschatzvermittlung. Frankfurt (M): Peter Lang.
- Houdé, O. (2004): Wie das kindliche Bewusstsein erwacht. In: Spektrum der Wissenschaft (Spezial 1): Bewusstsein, S. 32-36.
- Löschmann, M. (1993): Effiziente Wortschatzarbeit. Frankfurt (M): Peter Lang.
- Kühn, P. (2005): Mein Schul-Wörterbuch. Troisdorf: Wolf [und frühere Varianten].
- Kühn, P. (2007): Rezeptive und produktive Wortschatzkompetenzen. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Hohengehren: Schneider, S. 160-167.
- Miller, G. (1993): Wörter. Heidelberg: Spektrum.
- Nix, D. (2007): Vorleseübung als Wortschatztraining. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Hohengehren: Schneider, S. 157-158.
- Polenz, P. von (1988): Deutsche Satzsemantik. Berlin: de Gruyter.
- Lehmann, R. H./Peek, R./Gänsfuß, R./Husfeldt, V. (1996-2000): Aspekte der Lernausgangslage und der Lernentwicklung. Jahrgänge 5, 7, 9.
- Rudolph, G. (2007): „Diese Wörter kenne ich nicht.“ – Fachbegriffe in Sachtexten erschließen. In: Deutschunterricht, 60, 2, S. 20-25.
- Schlobinski, P. (1998): Jugendliche und „ihre“ Sprache: Sprachregister, Jugendkulturen und Wertesysteme. Opladen: Westdeutscher Verlag.
- Schmidt, L. (Hrsg.) (1973): Wortfeldforschung. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Schoonen, R./Hulstijn, J./Bossert, B. (1998): Metacognitive and Language-Specific Knowledge in Native and Foreign Language Reading Comprehension: An Empirical Study Among Dutch Students in Grades 6, 8 and 10. In: *Language Learning* 48, H. 1, S. 71-106.
- Ulrich, W. (2001): Didaktik der deutschen Sprache, Bd. 3 (Grammatik, Wortschatzarbeit, Unterrichtsmittel). Stuttgart: Ernst Klett.
- Ulrich, W. (2007): Wörter, Wörter, Wörter. Wortschatzarbeit im muttersprachlichen Deutschunterricht. Hohengehren: Schneider.
- Willenberg, H. (2007): Der vergessene Wortschatz. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Hohengehren: Schneider, S. 148-156.
- Willenberg, H. (2007): Wortschatz. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen: Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 130-139.

Michael Krelle / Heiner Willenberg

7 Argumentation Deutsch

7.1 Testkonstrukt

Mit dem Begriff Argumentation werden komplexe Anforderungen an Interaktanten in argumentativen Situationen beschrieben. In diesem Sinne zählt Argumentation auch zu den wesentlichen curricularen Anforderungen des Unterrichtsfaches Deutsch. So fordern die Lehrpläne unter anderem, dass Schüler mit Argumenten überzeugen und situationsadäquat zu einem Konsens kommen sollen.¹ Vor diesem Hintergrund untersucht die Studie Leistungen im argumentativen Bereich. Das Messmodell ist innerhalb der Argumentationstheorie und der Linguistik theoretisch verortet. Das Konstrukt umfasst sowohl kommunikativ-situative, als auch konzeptionell-reflexive Anforderungen des sprachlichen Handelns, die von je unterschiedlichen Wissenssystemen bestimmt werden.

Grundlegend wird davon ausgegangen, dass die argumentativen Anforderungen durch unterschiedliche Abstraktionsgrade der sprachlichen Tiefenstruktur darstellbar sind. Im Folgenden werden Grundkategorien des Modells angedeutet, die an anderer Stelle eine ausführlichere und theoretisch vertiefte Beschreibung finden (Willenberg/Gailberger/Krelle 2007; Krelle 2007). In der Argumentationsforschung haben kognitionswissenschaftliche Aspekte bisher wenig Beachtung gefunden. Diese können jedoch ein kohärentes Modell situativer und konzeptionell-reflexiver Anforderungen bieten (Hofer 2003). Aus dieser Perspektive stellen sich die Zusammenhänge von Positionen, Argumenten und Begründungen anhand unterschiedlich bewusster und abstrakter kognitiver Operationen der Interaktanten in Situationen dar. In der Kernstruktur stellen Argumentationen dabei die individuellen Erfahrungen von Interaktanten über die Zusammenhänge bestimmter Wissensbestände und -systeme in strittigen Situationen dar. Diese Dispositionen können als argumentative Konzepte sprachlicher Anforderungen bezeichnet werden. Zu den basalen Anforderungen zählen die Situationseinschätzung und die Auswahl des adäquaten sprachlichen Registers. Argumentative Handlungen erfordern ebenso das aktive Verfügen über ein eigenes inhaltliches Konzept des Begründens. Dabei ist es nötig, Aussagen mit subjektiven Aspekten kohärent zu ergänzen. Zeitgleich ist es erforderlich, die Aussagen von Interaktionspartnern zu verstehen, um dann sinnvoll und inhaltlich auf diese einzugehen. Zu den kognitiv anspruchsvollen Prozessen zählen die Reflexion von Argumenten und Strukturen, sowie deren bewusste Verwendung. Diese Anforderungen stellen sich für Interaktanten als unterschiedlich schwierige Abstraktionsstufen von der Textoberfläche dar, die curricular bereits in der Neunten Klasse gefordert werden.

¹ Es sind bei maskulinen Formen allgemeiner Bezeichnungen stets auch weibliche Personen mitgemeint.

7.2 Beschreibung der Kompetenzniveaus

Die Testergebnisse bestätigen das oben dargestellte Konstrukt, so dass aus testtheoretischen Gründen zwei Kompetenzniveaus modelliert werden. Diese entsprechen sowohl den fachwissenschaftlichen Annahmen, als dass sie auch wesentliche curriculare Ansprüche abbilden. Aus didaktischen Gründen wird das hier dargestellte Modell an anderer Stelle erweitert (Krelle/Vogt/Willenberg 2007), um es so für den Unterricht operabel zu machen.

Das Niveau A „Situation“ beschreibt die situativen und kommunikativen Aspekte. Zu diesen zählt, dass Interaktanten eine Situation in ihren Grundbedingungen und ihrer Gesprächsthematik einschätzen, an einem Gespräch teilnehmen und Aussagen auf ihre Tendenz hin situativ beurteilen können. In diesem Bereich sind ebenfalls Bedingungen der Argumentation anzusiedeln (Kopperschmidt 2000). Das Erkennen eines Dissens, das Verfügen über eine eigene Position und das Verfolgen eines Ziels gelten gleichermaßen als Voraussetzungen (Hofer 2003). Als argumentative Ziele werden zum einen das Markieren und das Vertreten der eigenen Position, der Wunsch nach Aufklärung einer Meinungsverschiedenheit und einer strittigen Frage gerechnet (Völzing 1979).

Das Niveau B „Reflexion“ fasst hingegen die konzeptionellen, strukturierenden Tätigkeiten zusammen. Aus dieser Perspektive stellen die Zusammenhänge von Positionen, Argumenten bzw. Begründungen unterschiedliche bewusste und abstrakte kognitive Operationen der Interaktanten in strittigen Situationen dar. Unter anderem erfordert Argumentation die Aktivierung unterschiedlich strukturierter Begründungen und deren Reflexion (Kopperschmidt 2006). Auch aus wissenschaftstheoretischer Perspektive beschreiben die Niveaus A und B zu unterscheidende Anforderungen. Für eine ausführlichere Diskussion der Ergebnisse verweisen wir ausdrücklich auf den von Heiner Willenberg (2007) herausgegebenen Band zur praktischen Anwendung der DESI Ergebnisse für den Deutschunterricht. Tabelle 7.1 gibt eine Übersicht über die beiden Kompetenzniveaus in Deutsch Argumentation.

Tabelle 7.1: Beschreibung der Kompetenzniveaus in Deutsch Argumentation.

Kompetenzniveau	Beschreibung
B (ab 717 Punkten)	Reflexion der argumentativen Struktur von Beiträgen oder Texten. Den argumentativen Aufbau eines Textes bzw. einer Rede alltags-sprachlich oder mit Fachbegriffen bezeichnen können.
A (486 bis 716 Punkte)	Einschätzung: Eine Situation in ihren Grundbedingungen und ihrer Gesprächsthematik einschätzen können. An einem Gespräch in der Situation teilnehmen können. Aussagen in der Situation beurteilen können.

Testaufgaben und Beispiele

Für die Testung von Argumentation gelten besondere Voraussetzungen und Gegebenheiten. Argumentative Konzepte werden gemeinhin anhand eines mehr oder weniger komplexen, kohärenten und intensionalen Beziehungsnetzes von Aussagen aus Gesprächsdaten oder aus Texten rekonstruiert. Dieses Beziehungsnetz lässt sich auch als sprachliches Oberflächenmuster kennzeichnen, anhand dessen die abstrakte Tiefenstruktur des Problemlösungskonzeptes ‚Argumentation‘ rekonstruiert werden kann. Zur Zeit der Testentwicklung (2002) waren nur sehr kostenintensive Verfahren zur Überprüfung videobasierter Gesprächsdaten mit großen Schülerinnen- und Schülerzahlen möglich. Aus diesem Grund rekonstruiert DESI die kognitiven Operationen nicht anhand von Sprachoberflächendaten aus Gesprächen, sondern führt die Schülerinnen und Schüler in fiktive schriftliche Situationen hinein, die dann selbstständig bewältigt werden müssen.

Tabelle 7.2: Aufgabenbeispiele für die beiden Kompetenzniveaus in Deutsch Argumentation.

Kompetenzniveau	Beispiele
A (486 bis 716 Punkte)	<p>Beispiel 1: geschlossenens Format</p> <p>Klaus hat die Sommerferien an einem See in den Alpen verbracht. Er ist dort zum ersten Mal Wasserski gefahren. Zu Hause trifft er seine Freunde Peter, Jochen und Armin und erzählt ihnen von seinen Erlebnissen. Dabei verfolgt Klaus zwei Ziele: er möchte nicht als Aufschneider auftreten, der angibt, aber er möchte wieder in die Gruppe reinkommen und bei seinen Freunden eine wichtige Rolle spielen.</p> <p>Was wird er sagen? Kreuze jeweils nur ein Kästchen an! (Erster Gesprächszug als Beispiel)</p> <p>Peter: „He, Jochen, guck mal, der Klaus ist wieder da!“ Jochen: „Prima, der kann gleich beim Bau der Höhle helfen.“ Armin: „Mannomann, ist der braun!“ Klaus:</p> <ul style="list-style-type: none"> a) „Ja, wir waren auch lange weg.“ b) „He, was macht ihr denn gerade?“ c) „Ach, es war nicht so doll.“ d) „Reisen mit „Orion“ machen es möglich!“ <p>(Bearbeitet nach: Franz Hebel 1976)</p>
	<p>Beispiel 2: offenes Format</p> <p>Begründe deine Wahl:</p> <hr/> <hr/> <hr/>
B (ab 717 Punkten)	

Wie oben angedeutet ergeben sich aus den Testergebnissen zwei Kompetenzniveaus, die im Folgenden anhand der Aufgabe „Wasserski“ exemplarisch vorgestellt werden. Bei dieser wird zunächst eine kommunikative Einschätzung gefordert, die dann im Gesprächsverlauf reflexiv erweitert werden muss. Hierfür ist es erforderlich, dass verschiedene Gesprächszüge der fiktiven Person „Klaus“ antizipiert werden. Bei den vorliegenden Beispielen, die in Tabelle 7.2 dargestellt sind, wird ein Gesprächszug in Auszügen abgebildet.

Die Aufgabe 1 erfordert bei der richtigen Antwortmöglichkeit b) adäquates situatives Wissen über die Kommunikationsgegebenheiten und integriert die Teilnehmer. Die Antwort c) drückt hingegen zu viel Bescheidenheit aus, Antwort a) und d) zielen zu stark auf die eigenen Position ab. Zu den Anforderungen des Niveaus B zählt hingegen das Formulieren einer Begründung für eine eingenommene Position. Hierfür müssen alltagssprachliche oder einfache pragmatische Begründungen vorgetragen werden.

Der Erstautor dieses Artikels entwickelt derzeit ein kostengünstiges diagnostisches Verfahren, um Argumentationen von Schülern curriculumsnah in realen Gesprächssituationen per Videoanalyse zu überprüfen (Krelle 2007).

Leistungsverteilung

Da die Datenerhebung des Tests aus den beschriebenen Kostengründen lediglich an einem Testzeitpunkt (am Ende der 9. Jahrgangsstufe) vorgenommen wurde, können keine Aussagen über Lernzuwächse getroffen werden. Die Prüfung der Leistungsverteilung auf die Kompetenzniveaus zeigt jedoch, dass das reflexive Niveau nur von wenigen Personen erreicht wird (1%). Hinzu kommt eine große Anzahl von Schülern unterhalb des geforderten Niveaus A (44%). Leistungen in diesem Bereich werden im Folgenden als „Unter Niveau A“ klassifiziert. Solche Schüler, die unterhalb dieser Erwartungsgrenze liegen, können die Hälfte der geforderten Aufgaben gelöst haben, es kann in solchen Fällen allerdings nicht von einer stabilen Leistungsdisposition gesprochen werden. Mehr als die Hälfte der Schüler erreichen hingegen das Kompetenzniveau A sicher (55%). Diese Schüler können konstant Situationen einschätzen und den Anforderungen des Sprechaktes gerecht werden.

Die Bedeutung der Bildungsgänge

Wie bereits angedeutet, bleiben viele Schüler unterhalb des einfachsten Niveaus. Besonders stark betrifft dieses Phänomen die Hauptschule (79%) und die IGS (65%). Auch an der Realschule erreichen immerhin noch mindestens ein Drittel nicht die formulierte Erwartungsgrenze (38%). Am Gymnasium hingegen tritt dieses Problem kaum zu Tage (10%), Gymnasialschüler lernen es, während ihrer Schullaufbahn zumeist, situativen Anforderungen adäquat zu begegnen (87%). In diesem Sinne scheinen verschiedene Lehrpläne eine hohe Erwartungshaltung zu formulie-

ren, die zumindest außerhalb der Gymnasien nicht eingehalten werden kann. Die Leistungsverteilung über die Bildungsgänge dokumentiert die Abbildung 7.1. Die senkrechten Markierungen weisen auf die Schwellen der Kompetenzskala hin, an denen das Niveau A oder B erreicht wird, während die Fläche der Kurve der jeweiligen relativen Zahl der Schülerinnen und Schüler entspricht.

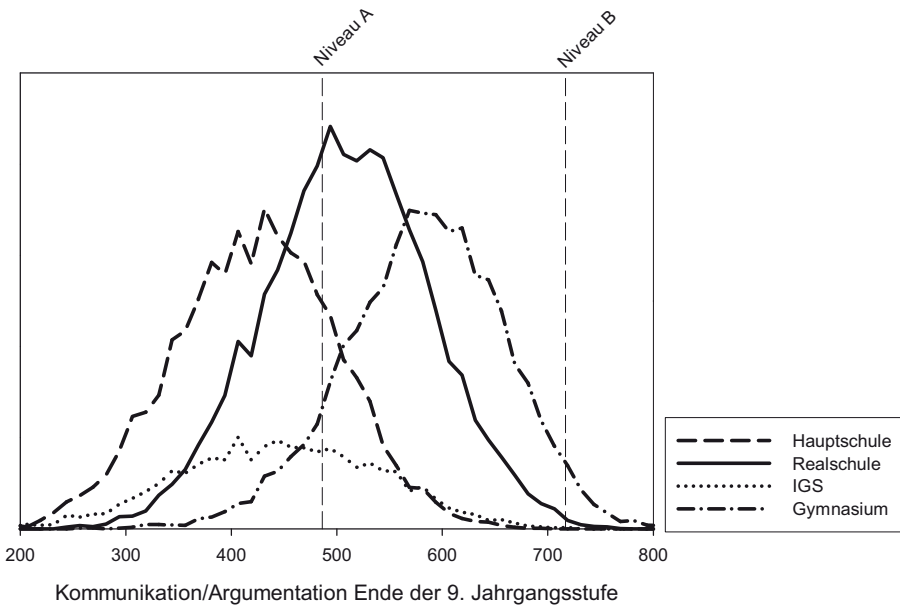


Abbildung 7.1: Leistungsverteilung Deutsch Argumentation nach Bildungsgängen.

Bei genauer Betrachtung der Leistungsverteilung lässt sich ein weiteres Phänomen konstatieren: Über die Hälfte der Schüler können zwar die situativen Anforderungen bewältigen, so liegen 55% auf dem Kompetenzniveau A. Bei der Strukturanalyse der Argumente und der Konzeptionalisierung der Argumentation haben hingegen fast alle Schülerinnen und Schüler erhebliche Schwierigkeiten, nur 3% am Gymnasium erreichen dieses anspruchsvolle Niveau sicher. An der Hauptschule und der IGS liegen die Ergebnisse eher unterhalb des Kompetenzniveaus A. Dennoch erreicht jeder dritte (35%) an der IGS bzw. jeder fünfte Schüler (21%) an der Hauptschule das Niveau A in diesem Bildungsgang. Lediglich solche Schüler lernen es, die curricularen Anforderungen zu bewältigen und stabil Situationen in ihren wechselseitigen Relationen einzuschätzen. Eine Erklärung dieser Polarisierung der Leistungsverteilung kann unter anderem anhand geschlechtsspezifischer und bezüglich der Erstsprache differenzierter Ergebnisse vorgenommen werden.

Die Bedeutung der Geschlechterdifferenz

Ein Vergleich der geschlechterspezifischen Daten zeigt, dass die Leistungen der Jungen im Mittel 61 Leistungspunkte unterhalb der Ergebnisse der Mädchen gleichen Alters liegen. In allen Bildungsgängen erreichen die Mädchen einen deut-

lich höheren Leistungsstand. Wenn man die Marke von 500 Leistungspunkten als Maßstab zu Grunde legt, können überdurchschnittliche Resultate der Mädchen gerade an der Realschule (537 Punkte) und am Gymnasium (605 Punkte) diagnostiziert werden. Auch wenn am Gymnasium die Jungen mit 556 Punkten überdurchschnittliche Leistungen erzielen, ist ebenfalls in diesem Bildungsgang ein deutlicher Vorsprung von 49 Leistungspunkten zugunsten der Mädchen zu konstatieren. Diese scheinen den curricularen Anforderungen mit Ausnahme der IGS besser zu genügen, als es bei den Jungen der gleichen Klassenstufe der Fall ist.

Die Bedeutung der Erstsprache

Überdurchschnittliche Ergebnisse im Gesamtest werden durch Schüler mit deutscher Erstsprache (512 Punkte) und mehrsprachiger Herkunft erzielt (503 Punkte). Allerdings bewältigen auch Schüler ohne deutsche Erstsprache am Gymnasium mit 545 Leistungspunkten die Anforderungen sehr gut. In allen weiteren Bildungsgängen erbringen Schüler ohne Deutsch als Erstsprache schlechtere Ergebnisse (z.B. IGS 400 Leistungspunkte). Dieses lässt darauf schließen, dass neben der Beherrschung der Deutschen Sprache auch soziokulturelle Faktoren einen entscheidenden Einfluss auf die Leistungswerte in diesem Bereich haben (vgl. auch Kapitel 6). Hinweise dafür sind ebenfalls die Leistungen von Schülern ohne Erstsprache Deutsch an der Hauptschule (401 Punkte), der IGS (400 Punkte) und an der Realschule (478 Punkte), die allesamt klar unter dem genannten Maßstab von 500 Punkten liegen. In diesem Zusammenhang muss konstatiert werden, dass die Realschüler nicht deutscher Erstsprache (478 Punkten) über den Leistungen von Schülern mit deutscher Erstsprache an der Hauptschule (429) und der IGS (466) rangieren.

Betrachtet man die Leistungs differenzen in Bezug auf die sprachliche Herkunft, so verstärkt sich das bisherige Bild. Die größten Unterschiede können zwischen Schülerinnen und Schüler mit und ohne deutsche Erstsprache gemessen werden. Insbesondere an der IGS ist diese Differenz mit 67 Punkten ausgeprägt. Da auch der Unterschied der Erstsprachler Deutsch zu den mehrsprachig aufgewachsenen Schülern mit 38 Punkten Differenz hoch ist, scheinen die Schülerinnen und Schüler der IGS am stärksten von einer Sprachbarrierenproblematik betroffen zu sein und den curricularen Anforderungen in diesem Bereich nur bedingt zu genügen. In dieser Ausprägung tritt das Phänomen in keinem anderen Bildungsgang auf.

Implikationen für die Deutschdidaktik und den Deutschunterricht

Der curriculare Anspruch im Bereich Argumentation scheint gymnasial bestimmt zu sein und nicht mit den Kompetenzen von Schülern in den anderen Bildungsgängen der neunten Klasse übereinzustimmen. Die Bildungsstandards und ländereigenen Lehrpläne beanspruchen von „hoher Warte“ aus, was Schüler nur sehr eingeschränkt einlösen können. Es ist allerdings nötig, Argumentation als eine der wichtigsten Kommunikationsformen überhaupt einzusetzen und anstelle der Auseinandersetzung

mit bloßen Behauptungen die diskursive und argumentative Form wählen zu können.

Interessante Ergebnisse in dieser Hinsicht zeigt das Forschungsprojekt von Grundler und Vogt (2005), bei dem gleichsam habituelle Differenzen des Argumentierens beobachtet wurden: Hauptschüler neigen zur bloßen Konstatierung, Gymnasiasten suchen hingegen kohärente Begründungen. In diesem Sinne müssen die Probleme in Bezug auf die Bildungsgänge, die Geschlechterdifferenzen und die Sprachproblematik auch curricular angegangen werden. Offenkundig sollte kohärente Argumentation zunächst in seinen einfacheren Varianten geübt werden. Da kommunikative Teiltrainings die Argumentation vorbereiten, bieten sich aus der Sicht von DESI holistische Übungen zur Situationseinschätzung an. Darauf aufbauend kann u.a. die Beschäftigung mit transkribierten Gesprächen stehen (Vogt 2004; Krelle/Vogt/Willenberg 2007). Für weitergehende Empfehlungen bezüglich der praktischen Anwendungen der Ergebnisse verweisen wir auf noch folgende Veröffentlichungen beider Autoren.

Das reflexive Niveau wird nur von wenigen Schülerinnen und Schüler erreicht. Dieser Effekt verwundert zunächst, da die deutschdidaktischen Vorgaben gerade solche Anforderungen zu fördern versuchen. Der Erstautor entwickelt aus diesem Grund einen Forschungsansatz, der Argumentation in Gesprächssituationen an der ihr eigenen Charakteristik fasst – der Aushandlung konkurrierender Wissens Elemente durch Personen. Gerade die kognitiv anspruchsvolleren Tätigkeiten erfordern die Aktivierung von Konzepten und Wissenssystemen. Diesen Wissenssystemen kommt ein situativer Geltungsstatus zu, der durch verbale Interaktion geklärt werden soll. Als empirisch-didaktischer Ansatz stellt die „Wissensbasierte Argumentation“ situative und konzeptionell-reflexive Anforderungen auf eine systematische Grundlage und ermöglicht so, soziokulturell oder sprachlich benachteiligte Schüler zur diskursiven Problemlösung anzuregen.

Literatur

- Brinker, K./Sager, S.F. (2001): Linguistische Gesprächsanalyse. Berlin: Wilhelm Schmidt.
- Deppermann, A./Hartung, M. (Hrsg.) (2003): Argumentieren in Gesprächen. Tübingen: Stauffenburg.
- Grundler, E./Vogt, R. (Hrsg.) (2005): Argumentieren in Schule und Hochschule. Tübingen: Stauffenburg.
- Hofer, Manfred (2003) Argumentieren. In: Herrmann, Th./Grabowski, J. (Hrsg.): Enzyklopädie der Psychologie: Sprachproduktion. Göttingen: Hogrefe, S. 801-824.
- Kienpointner, M. (1992): Alltagslogik. Stuttgart-Bad Canstatt: Frommann-Holzboog.
- Kintsch, W. (1988): The Role of knowledge in discourse comprehension: A construction-integration model. In: Psychological Review 95, S. 163-182.
- Kintsch, W. (1998): Comprehension. A paradigm for cognition. New York: Cambridge University Press.
- Klein, W. (1981): Logik der Argumentation. In: Schröder, P./Steger, H. (Hrsg.): Dialogforschung. Jahrbuch 1980 des Instituts für deutsche Sprache, Düsseldorf: Pädagogischer Verlag Schwann, S. 226-264.

- Kopperschmidt, Josef (2000): Argumentationstheorie zur Einführung. Hamburg: Junius.
- Kopperschmidt, Josef (2006): Argumentationsarbeit. Oder: Über die Arbeit mit der goldenen Kette. In: Grundler, E. & Vogt, R. (Hrsg.): Argumentieren in Schule und Hochschule. Tübingen: Stauffenburg, S. 11-28.
- Kindt, Walther (1999): Was sollte man in der Schule über Argumentationen lernen, „Rhetorik und Argumentation“. In: Der Deutschunterricht 5, S. 26-36.
- Krelle, M. (2007): Wissensbasierte Argumentation lehren und lernen. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Hohengehren: Schneider, S. 108-114.
- Krelle, M./Vogt, R./Willenberg, H. (2007): Argumentative Kompetenz im Mündlichen. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Hohengehren: Schneider, S. 96-107.
- Strohner, H. (2000): Kognitive Voraussetzungen: Wissenssysteme - Wissensstrukturen - Gedächtnis. In: Brinker u.a. (Hrsg.): Text- und Gesprächslinguistik. Ein internationales Handbuch. New York: De Gruyter, S. 261-274.
- Strohner, H./Brose, H. (2000): Die Rolle von Wissenssystemen für die Gestaltung interaktiven Handelns. In: Brinker u.a. (Hrsg.): Text- und Gesprächslinguistik. Ein internationales Handbuch. New York: De Gruyter, S. 1169-1177.
- Strohner, H. (Hrsg.)/Brose, H. (Mitarb.) (2001): Kommunikation. Kognitive Grundlagen und praktische Anwendungen. Wiesbaden: Westdeutscher Verlag.
- Vogt, R. (2006): Argumentieren multimodal. Perfromanzlinguistische Perspektiven. In: Grundler, E./Vogt, R. (2006): Argumentieren in Schule und Hochschule. Interdisziplinäre Studien. Tübingen: Stauffenburg Linguistik, S. 41-62.
- Vogt, R. (2002): Im Deutschunterricht diskutieren. Zur Linguistik und Didaktik einer kommunikativen Praktik. Tübingen: Niemeyer.
- Vogt, R. (2004): „Miteinander-Sprechen lernen: Schulische Förderung von Gesprächsfähigkeit“. In: Becker-Mrotzek, M./Brünner, G. (Hrsg.): Analyse und Vermittlung von Gesprächskompetenz. Frankfurt: Peter Lang, S. 199-224.
- Vözing, P.-L. (1979): Begründen, Erklären, Argumentieren: Modelle und Material zu einer Theorie der Metakommunikation. Heidelberg: Quelle und Meyer.
- Willenberg, H./Gailberger, S./Krelle, M. (2007): Argumentieren. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Testinstrumente zur Messung der Leistung im Deutschen und Englischen. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Willenberg, H. (Hrsg.) (2007): Kompetenzhandbuch für den Deutschunterricht. Hohengehren: Schneider.

Astrid Neumann / Rainer H. Lehmann

8 Schreiben Deutsch

8.1 Testkonstrukt

Das DESI-Konstrukt Textproduktion ist in einer dreifachen theoretischen Verortung innerhalb der linguistischen Text-, Schreibprozess- und Schreibentwicklungsfor-schung, der Curriculumforschung und der empirischen Aufsatzforschung entstanden (vgl. Harsch u.a. 2007). Schreibfähigkeit wird dabei im Sinne der funktio-nalen Textlinguistik (Brinker 1988; deBeaugrande/Dressler 1981; Nussbaumer 1991) verstanden, bei der Texte als „komplex strukturierte, thematisch wie konzeptuell zusammenhängende sprachliche Einheit[en] verstanden werden, A. N.], mit der ein Sprecher eine sprachliche Handlung mit erkennbarem kommunikativen Sinn voll-zieht“ (Linke/Nussbaumer/Portmann 1996, S. 245). Daraus leitet sich die folgende Definition ab:

„Unter Schreibkompetenz¹ wird hier die Fähigkeit verstanden, Texte adres-satengerecht zu formulieren und, je nach Zielsetzung, präzise zu informieren, überzeugend zu argumentieren oder Sprache ästhetisch ansprechend und kre-ativ einzusetzen“ (Harsch u.a. 2007, S. 45).

Die Schülerinnen und Schüler erhielten die Möglichkeit, anhand von authentischen Schreibsituationen, dem Schreiben von Briefen, diese Fähigkeit zu zeigen. Uns ist bewusst, dass in einer Testsituation – einer künstlich geschaffenen –, wie sie bei der Entstehung der Schreiben des DESI-Aufsatzkorpus zugrunde lag, eventuell Verzerrungen im Sinne von Kompetenz-Performanz-Modellen entstehen, da sich die Schreibenden in „realen Schreibsituationen“ anders verhalten könnten. Daher sind die gezeigten Leistungen als *Indikatoren* für dahinter liegende Schreibkompetenzen bzw. deren Defizite zu verstehen.

Da der Auftraggeber (KMK) die Untersuchung von Schülerleistungen in der neunten Jahrgangsstufe anhand von Texten forderte – nicht eine Beschreibung der Lernentwicklung bis zum Testzeitpunkt und auch nicht die Beschreibung des Schreibprozesses in der Testsituation –, ist es primäres Untersuchungsziel des Tests, konkrete Aussagen über die in den geschriebenen Texten gezeigten Leistungen der Schülerinnen und Schüler zu treffen. In dem Maße, wie die Daten es zulassen, kön-nen in späteren Analysen Topoi der oben genannten Bezugsdisziplinen aufgegriffen werden.

1 Schreibfähigkeit und Schreibkompetenz werden hier synonym verwendet.

8.2 Testaufgaben

Aufgrund einer umfangreichen Curriculum-Analyse wurden vier verschiedene, mit zwei inhaltlichen und zwei formellen Schwerpunkten weitgehend parallelisierte Briefe als Schreibaufgaben entwickelt, die in einem rotierten Design über beide Messzeitpunkte in der neunten Jahrgangsstufe zu bearbeiten waren. Es handelte sich dabei um

- ein Reklamationsschreiben wegen einer fehlerhaften Computerlieferung;
- einen Brief an eine Freundin/einen Freund wegen einer fehlerhaften Computerlieferung;
- einen Beschwerdebrief wegen der Schließung eines Jugendclubs;
- einen Brief an eine Freundin/einen Freund wegen der Schließung eines Jugendclubs.

Dabei erhielt jede Schülerin/jeder Schüler zu einem gegebenen Messzeitpunkt in der neunten Jahrgangsstufe jeweils eine Aufgabe aus den beiden inhaltlichen Bereichen, so dass Erinnerungseffekte ausgeschlossen werden konnten. Auf diese Weise wurde eine systematische Berücksichtigung der spezifischen Aufgabenschwierigkeiten in der Feststellung der Schülerleistungen und in der Beobachtung von Schreibentwicklungen innerhalb der neunten Jahrgangsstufe ermöglicht.

Für den ersten Messzeitpunkt zu Beginn der neunten Jahrgangsstufe liegen 10056 Schreibprodukte vor, für den zweiten Messzeitpunkt am Ende der neunten Jahrgangsstufe 9844. Der Anteil nicht auswertbarer Arbeiten² stieg vom ersten zum zweiten Messzeitpunkt in der neunten Jahrgangsstufe von 2% auf 4%.

8.3 Kodierung

Jeder Text wurde im Doppel-Blind-Verfahren von zwei unabhängigen, intensiv geschulten Ratern kodiert. Dabei wurden zwischen 43 und 47 dichotome inhaltliche und formelle Merkmale und 6 Ratingkategorien zu textsortenspezifischen sowie sprachsystematischen Merkmalen erfasst. Die Beurteilerübereinstimmung bei den dichotom zu kodierenden Merkmalen ist mit insgesamt 96% zu beiden Messzeitpunkten in der neunten Jahrgangsstufe sehr zufriedenstellend. Für die fünfstufigen Ratingskalen konnten absolute Übereinstimmungen der Beurteilung für 56% bzw. 59% erreicht werden, ergänzt man diese Einschätzungen um diejenigen, in denen sich die Bewertungen beider Rater um eine Stufe unterschieden, so sind damit 95% bzw. 96% erfasst. Die Reliabilitäten der Ratingskalen liegen mit einem durchschnittlichen α von .82 zum ersten Messzeitpunkt (Beginn der neunten Jahrgangsstufe) bzw. von .84 zum zweiten (Ende der neunten Jahrgangsstufe) in einem für die Einschätzung ersprachlicher Kompetenzen hoch akzeptablen Bereich (vgl. Wirtz/Caspar 2002). Trotzdem gibt es bei allen Textrezeptionsprozessen einen verbleibenden Anteil an raterspezi-

2 Darunter fallen nicht ausgefüllte Testblätter und Texte, die zu anderen Themen (Bsp. eine Bestellung oder Hinweise an das Testteam) geschrieben wurden und damit nicht valide im Sinne der Testung sind (s. Neumann 2007a).

fischen Besonderheiten in der Einschätzung der Schülertexte sowohl innerhalb der Texte eines Messzeitpunktes in der neunten Jahrgangsstufe als auch zwischen diesen beiden.³ Die verschiedenen Ausprägungen der Milde-Strenge-Tendenz wurden statistisch durch die Berücksichtigung von Ratereffekten in der Rasch-Skalierung der Daten über beide Messzeitpunkte in der neunten Jahrgangsstufe korrigiert (*Facetten-Modelle*), sodass von einer objektiven Erfassung der Schülerleistungen im Schreiben ausgegangen werden kann.

8.4 Skalenbildung

Analysen mit den Daten der Textproduktion in Deutsch haben die Befunde der IEA-Aufsatzstudie (Lehmann 1988, 1990) repliziert und gezeigt, dass es sich beim Schreiben um ein zweidimensionales Konstrukt handelt, bei dem für die Erstellung eines Textes zum einen semantisch-pragmatische Komponenten beherrscht, zum anderen aber sprachsystematische Fähigkeiten aktiviert werden müssen. Beide Teilkompetenzen korrelieren mit $r = .56$ zu Beginn der neunten Jahrgangsstufe und $r = .62$ am Ende der neunten Jahrgangsstufe nur in einer mittleren Größenordnung. Daher wird hier für das Modul Textproduktion nicht ein kombinierter Testwert berichtet, sondern es werden für jede der Skalen „Semantik/Pragmatik“ und „Sprachsystematik“ separate Analysen vorgelegt.

Darüber hinaus konnte gezeigt werden, dass die inhaltlichen Komponenten der Texte, die hier mit dichotomen Merkmalen erfasst wurden, nicht unabhängig von ihrer semantisch-pragmatischen Realisierung sind (vgl. Neumann 2007b). Zugunsten der Berücksichtigung der Raterspezifika sind diese (aufgrund des anderen Datenformates) nicht in die Skalierungen eingegangen. Sie können aber wegen einer separaten Skalierung der Einzelitems zu einer Inhaltsskala deskriptiv in ihren spezifischen Anforderungen beschrieben werden, sodass im Detail etwas darüber ausgesagt werden kann, *was* die Schülerinnen und Schüler schreiben und *wo* besondere Stärken und Schwächen der inhaltlichen Arbeit liegen.

Neben der oben erwähnten Berücksichtigung der Milde- bzw. Strenge-Tendenz der Rater sind bei der Skalierung die Spezifika der Aufgaben selbst berücksichtigt worden, um eine Benachteiligung derjenigen Schülerinnen und Schüler auszuschließen, die relativ gesehen schwerere Aufgaben zu bewältigen hatten. Die Ergebnisse dafür sind in Tabelle 8.1 dargestellt.

3 So konnte anhand einer Replikationsstudie von 800 Texten des ersten Messzeitpunktes zu Beginn der neunten Jahrgangsstufe nach Beendigung der Bewertungen des zweiten Messzeitpunktes am Ende der neunten Jahrgangsstufe nachgewiesen werden, dass sich die Rater bei allen Skalen außer Rechtschreibung und Grammatik im Laufe des Prozesses durch eine nominell moderatere Bewertung in ihrer Bewertungspraxis tendenziell an die durchschnittlich schwachen Leistungen anpassen. Andererseits bewerten die Rater die Kategorien Rechtschreibung und Grammatik im Laufe des Prozesses nominell strenger, was auf eine differenziertere Sicht auf die typischen Fehler durch fortlaufende intensive Schulungen hindeutet. Somit wird die Notwendigkeit der Korrektur der Raterspezifika deutlich, damit die reinen Lerneffekte nicht durch abweichende Bewertungstendenzen überlagert sind (Neumann 2006).

Tabelle 8.1: Aufgabenspezifische Schwierigkeiten für die Skalen „Semantik/Pragmatik“ und „Systematik“ in Punkten auf den jeweiligen DESI-Skalen.

	Semantik/Pragmatik	Sprachsystematik
Computer formal	498	477
Computer persönlich	512	465
Jugendclub formal	529	495
Jugendclub persönlich	479	479

Es zeigen sich bei der Skala „Semantik/Pragmatik“ keine systematischen textsorten- oder themenspezifischen Besonderheiten. Als die leichtesten Texte erwiesen sich hier persönliche Schreiben zum Jugendclubthema, während die schwersten die zum selben Thema formal zu gestaltenden Texte sind. Dagegen sind beide Texte zum Computertema relativ gesehen durchschnittlich schwer, wobei sich hier die formellen Schreiben als minimal leichter erweisen. Die Schwierigkeit einer Aufgabe hinsichtlich der semantisch-pragmatischen Anforderungen hängt demnach nicht systematisch vom Thema oder der geforderten Textsorte, sondern von der spezifischen Aufgabenformulierung ab.

Anders dagegen verhält es sich bei der Skala „Sprachsystematik“, bei der themenabhängige schwierigkeitsdifferenzierte Schreibungen zu erkennen sind, wenn auch nur in geringem Umfang. Für die Briefe zum Thema „Computer“ ist die normgerechte Abfassung für die Schülerinnen und Schüler systematisch leichter gewesen als für die Reaktionen auf das Problem der Schließung des Jugendclubs. Hier scheint es den Schreibenden durch die Verwendung eines spezifischen Wortschatzes leichter zu fallen, sich schriftlich normgerecht auszudrücken. Darüber hinaus ist auffällig, dass in beiden Fällen jeweils die persönlichen Schreiben, in denen die Schreibenden eher auf die ihrem Sprechstil entsprechenden Formulierungen zurückgreifen können, relativ gesehen leichter zu bewältigen sind, als die formellen Schreiben, in denen auf formalisierte, eher ungebräuchliche Wörter und Wendungen zurückgegriffen werden muss.

8.5 Beschreibung der Kompetenzniveaus⁴

Skala „Semantik/Pragmatik“

Die für die Schreibenden am leichtesten zu erfüllende Kategorie innerhalb dieses Konstruktes ist die Stilkomponente mit einer spezifischen Schwierigkeit von 485 Punkten, gefolgt von der Wortwahl (511 Punkte), am schwersten ist die strukturell gebundene Komponente Textaufbau (517 Punkte). Es fällt den getesteten Schülerinnen und Schülern offenbar, und dies unabhängig von der zu erfüllenden Aufgabe, tendenziell schwerer, in einer angemessenen Wortwahl einen gut strukturierten, logisch aufgebauten Text zu verfassen. Die stilistischen Anforderungen der Textsorte bzw. des Genres mit dem entsprechenden Adressatenbezug und die Berücksichtigungen

⁴ Vgl. Harsch u.a. (2007).

für das Gelingen der Kommunikation dagegen fallen ihnen vergleichsweise leichter. Insgesamt ergeben sich in dieser Teildimension drei Niveaus, die in Tabelle 8.2 erläutert sind.

Tabelle 8.2: Beschreibung der Kompetenzniveaus in Textproduktion Deutsch: Semantik/ Pragmatik.

Kompetenz-niveau	Beschreibung
> C (ab 768 Punkte)	Sprachliche stilistisch-pragmatische Ausführung besser als C, ohne Inkonsistenzen.
C (613 bis 767 Punkte)	In sich schlüssiger, konsequent logisch aufgebauter Text, der alle wesentlichen Textelemente enthält; mit eventuell vorhandener Absatzgliederung zur Unterstützung der Argumentation, in der entsprechend der Briefform geforderten Stilistik (offiziell vs. persönlich) und Adressatenorientierung mit treffsicherer, zielführender, abwechslungsreicher Wortwahl; gelegentliche Inkonsistenzen können noch auftreten.
B (445 bis 612 Punkte)	Meist in sich schlüssiger, logisch folgerichtig aufgebauter Text, in dem die wesentlichen Textelemente zumindest in Ansätzen zu erkennen sind und der evtl. sich widersprechende Aussagen und Ansätze zu einer Absatzgliederung enthält; in weitgehend von der geforderten Briefform entsprechender Stilistik (offiziell vs. persönlich) und mit einer Adressatenorientierung mit angemessenem Ausdruck von Emotionalität; eigenes, allerdings begrenztes Wortmaterial wird im Wesentlichen sicher und treffend verwendet bei gelegentlichen Wortwiederholungen und Übergeneralisierungen.
A (340 bis 444 Punkte)	Kurzer Text, bei dem die Intention, die Aufgabe zu lösen, erkennbar ist, jedoch mit Mängeln, die eine reale Kommunikation belasten würden: Logisch fehlerhaft aufgebauter Text, wesentliche Textelemente eines Briefes fehlen, ohne Absatzgliederung, ohne bzw. mit fehlerhafter Adressatenorientierung, stilistisch der angestrebten Briefform (offiziell vs. persönlich) nicht entsprechend, mit sehr begrenztem Wortmaterial, das vorrangig dem Schreibimpuls entnommen ist, mit häufigen Wortwiederholungen oder unangemessenen Wörtern und Wendungen.

Betrachtet man die Ausführung aller inhaltlichen und formalen Merkmale durch die Schreibenden über alle Aufgaben zusammen (vgl. Abbildung 8.1), so lassen sich mit Hilfe der spezifischen Itemschwierigkeiten Aussagen darüber treffen, was den Schülerinnen und Schülern relativ leicht (< 500 Punkte) und was ihnen besonders schwer fiel (> 500 Punkte).

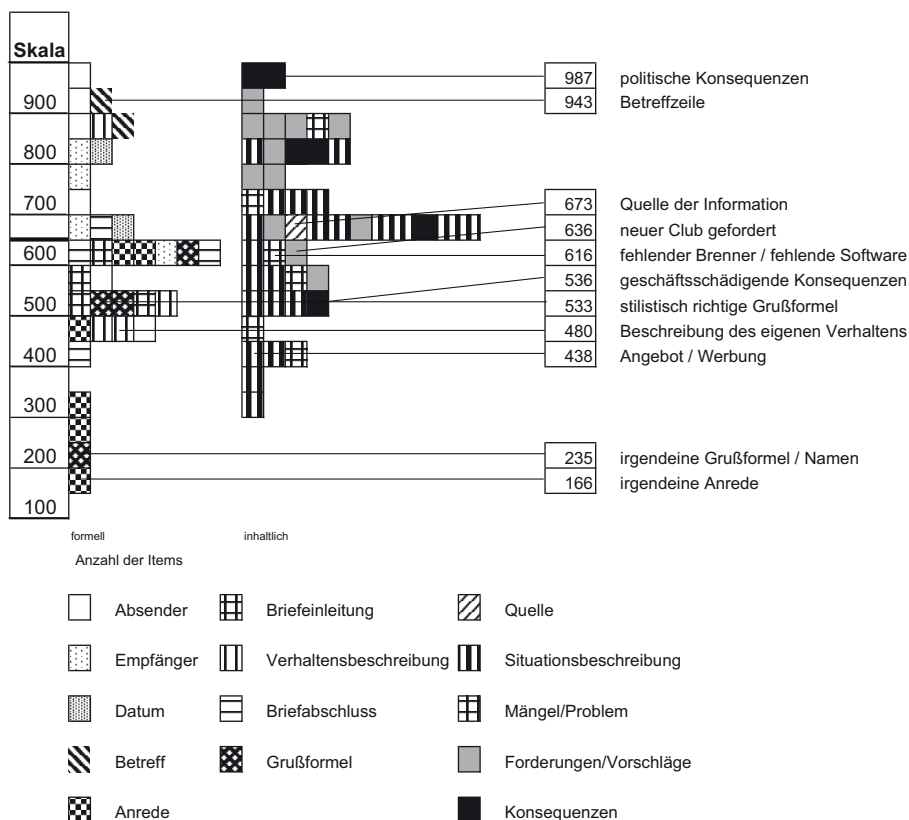


Abbildung 8.1: Verteilung der Einzelitems der Skala „Inhaltliche und formale Merkmale“ zum zweiten Messzeitpunkt am Ende der neunten Jahrgangsstufe.

Insgesamt kann festgestellt werden, dass den Schülerinnen und Schülern das Briefformat mit Anrede und Grußformel bekannt ist, dass es aber weniger Sicherheit in der Verwendung der stilistisch angemessenen Varianten derselben gibt. Auch Elemente wie die Briefeinführung und eine Abschlussequenz liegen ungefähr im selben Schwierigkeitsbereich wie die stilistisch richtigen Komponenten der Anrede. Datumsangaben zur Orientierung des Empfängers dagegen gehören bereits zu den schwereren Anforderungen an die Schreiben.

Im inhaltlichen Bereich fällt auf, dass Beschreibungen der Situation als Verstehensvoraussetzungen für den Empfänger zu großen Teilen bereits in den mittelschweren Bereich fallen. Besonders leicht sind das Aufzählen der Mängel und Probleme, sehr schwer fallen Lösungsansätze und mögliche Konsequenzen für den Empfänger. Die inhaltliche Umsetzung der inneren Logik der Schreiben, die in jedem Fall auf das Erreichen eines Ziels gerichtet ist, fällt den Schreibenden besonders schwer und wird auch nur von relativ wenigen Schülerinnen und Schülern gemeistert.

Um die Anforderungen an die Schreibenden genauer zu spezifizieren, wurden mehrere, für die Schwierigkeit der Aufgabenerfüllung verantwortliche Merkmale identi-

fiziert, die in Anknüpfung an theoretische Annahmen eines von der kognitiven und der Schriftsprachentwicklung dominierten Entwicklungsmodells zum kompetenten Schreiber abgeleitet wurden (genauer bei Neumann 2007). Danach unterscheiden sich die Lösungen der Einzelitems hinsichtlich der thematischen Ausrichtung (technisch vs. sozial; individuell vs. gruppenspezifisch), des Stimulus (graphisch vs. textuell; Hinweise der Aufgabenstellung), der Sprachebene (schriftsprachlich-reglementiert vs. Nähe zur konzeptionellen Mündlichkeit⁵), der Versprachlichung (aus Stimulus/Aufgabenstellung Bekanntes vs. Einbringen eigener Erfahrungen; Verwendung von Details) und der kommunikativen Sprachhandlung (Empfängerorientierung, Darstellung der Problemlage, Lösungsorientierung, persönliche Bezüge, Darstellung eigener Emotionen).

Mittels einer multiplen Regressionsanalyse der spezifischen Itemschwierigkeiten der Skala „Inhaltliche und formelle Merkmale“ auf diese schwierigkeitsbestimmenden Merkmale (vgl. Verfahren bei Hartig 2007) konnten ca. 50% der Schwierigkeitsunterschiede vorhergesagt werden. Durch eine Verankerung der Merkmale über die erwarteten Aufgabenschwierigkeiten wurden drei Anforderungskategorien bestimmt, die die inhaltliche Füllung der semantisch-pragmatischen Komponente wiedergeben (vgl. Tabelle 8.3).

Tabelle 8.3: Beschreibung der inhaltlichen Anforderungskategorien der Kompetenzniveaus in Textproduktion Deutsch: Semantik/Pragmatik.

Anforderungskategorie	Beschreibung
> C	Eigenständige Aufgabenbewältigung mit differenzierter Darstellung der Problemlage und Entwicklung von Lösungen und Konsequenzen; allgemeine Normen der Schriftsprache
B/C*	Extraktion von im Text präsentierten Einzelinformationen; textsorten-spezifische Normen; Emotionalität; gruppenspezifischer Rahmen
A	Extraktion von graphisch oder in der Aufgabenstellung präsentierten Einzelinformationen; Normen der mündlichen Umgangssprache; individueller Rahmen

Anmerkungen: *B und C unterscheiden sich hier nur durch die stilistisch-textsortenorientierte (das **Wie** der) Ausführung.

Skala „Sprachsystematik“

Die geringsten Anforderungen innerhalb des Konstruktes Sprachsystematik stellt die Komponente „Grammatik“ (392 Punkte), wobei darunter ausschließlich morphologische und morphematische Gegebenheiten kodiert wurden. Die Schülerinnen und Schüler sind in diesem Teilbereich relativ zu den anderen beiden in der neunten Jahrgangsstufe sicherer. Schon die Satzkonstruktion (505 Punkte) dagegen ist über

5 Für unsere Einschätzung sind dabei vor allem die Form der Sprache und die „Kommunikative Grundhaltung“ (vgl. Nussbaumer 1991, S. 270 ff) entscheidend, wobei zu beachten ist, dass diese verwendete Gegenüberstellung von Mündlichkeit und Schriftlichkeit einen idealtypischen Zustand darstellt, der durch graduelle Abstufungen repräsentiert wird.

alle Aufgaben hinweg schwieriger zu bewerkstelligen, die Schülerinnen und Schüler formulieren den Texten weniger angemessene Sätze oder verlieren sich in hochkomplexen Satzstrukturen, welche die Verständlichkeit der Schreiben negativ beeinflussen. Noch problematischer sind die spezifischen Fähigkeiten in der Komponente Rechtschreiben (540 Punkte). Geht man davon aus, dass die Schreibenden in freien Texten Wörter und Wendungen verwenden, die sie – anders als bei einem Diktat – selbst wählen, dann ist die (Nicht-)Beherrschung von normgerechtem, sprachsystematischem Schreiben hier besonders durch Defizite in der Orthographie bestimmt. Dies zeigt sich auch in der in Tabelle 8.4 aufgelisteten Beschreibung der drei Niveaus.

Tabelle 8.4: Beschreibung der Kompetenzniveaus in „Textproduktion Deutsch: Sprachsystematik“.

Kompetenz-niveau	Beschreibung
> C (ab 754 Punkten)	Fehlerfreier Text in variablem, abwechslungsreichem, angemessen komplexem Satzbau.
C (642 bis 752 Punkte)	Der Text entspricht den orthographischen Konventionen bei gelegentlich auftretenden Flüchtigkeitsfehlern. Die Wortgrammatik ist fehlerfrei. Es werden abwechslungsreiche, angemessen komplexe Satzkonstruktionen gewählt, die nur durch gelegentliche Inkonsistenzen in der Interpunktion gekennzeichnet sind.
B (471 bis 641 Punkte)	Der Text entspricht weitgehend den orthographischen Konventionen, es treten Probleme bei Fremd- und Fachwortschatz und komplexen Kombinationen auf. Die Wortgrammatik ist durch gelegentliche Inkonsistenzen, die die Verständlichkeit nicht beeinflussen, gekennzeichnet. Es werden angemessen komplexe Satzkonstruktionen verwendet, eine meist noch fehlerhafte Interpunktion erschwert aber den Lesefluss.
A (305 bis 470 Punkte)	Der Text ist bis auf einzelne, isolierte Passagen verständlich, Fehler in der Orthographie, Zeichensetzung und Satzkonstruktion wirken sich kommunikationsbelastend aus. Wortgrammatische Konventionen werden außer bei Komplexitätsproblemen weitgehend eingehalten.

8.6 Ergebnisse

Leistungsverteilung am Ende der neunten Jahrgangsstufe

Nachdem in den oberen Ausführungen die Konstrukte und ihre spezifischen Anforderungen vorgestellt wurden, sollen jetzt die Leistungen der Schülerinnen und Schüler in der neunten Jahrgangsstufe im Modul Textproduktion Deutsch in den genannten beiden Teildimensionen beschrieben werden.

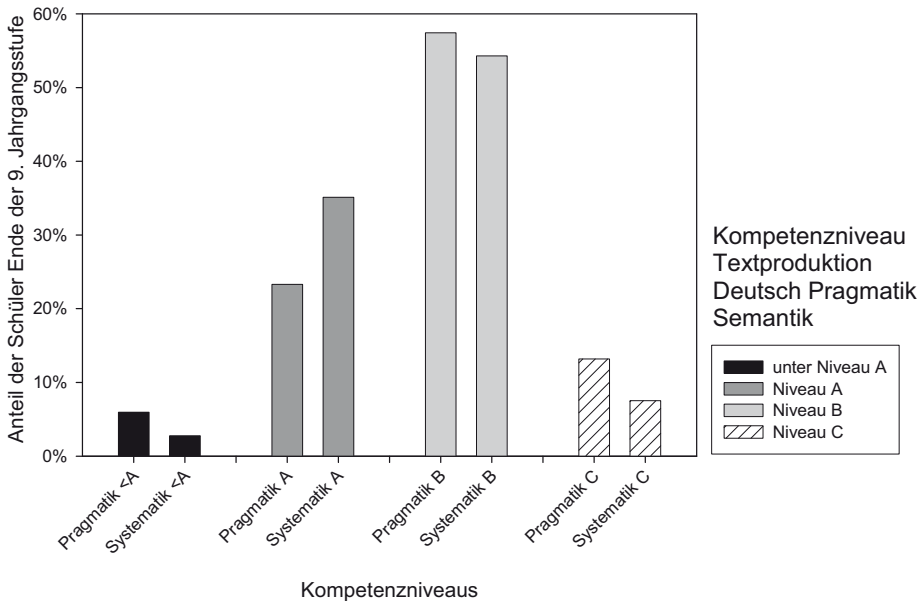


Abbildung 8.2: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Textproduktion Deutsch: Semantik/Pragmatik“ und „Systematik“ am Ende der neunten Jahrgangsstufe.

Wie in Abbildung 8.2 dargestellt, verfügen mehr als zwei Drittel (71%) der deutschen Schülerinnen und Schüler der neunten Jahrgangsstufe über eine semantisch-pragmatische Schreibkompetenz des Niveaus B und darüber. Diese Schreibenden besitzen die Fähigkeit, einen Brief mehr oder weniger adressatengerecht zu formulieren und den Empfänger über ihr Problem zu informieren. In stilistisch angemessener Art und Weise haben dies in dieser Alterstufe etwa 13% bewältigt, wobei sogar von 0.1% der Schreibenden „meisterliche“ Briefe vorgelegt wurden. Für den sprachsystematischen Bereich lassen sich ähnliche Befunde aufzeigen. Etwa zwei Drittel der Schülerinnen und Schüler (62%) schreiben auf einem angemessenen Niveau, das eine problemlose Kommunikation ermöglicht, davon sind allerdings nur 0.3% fehlerfreie Texte; 8% sind durch leichte Inkonsistenzen gekennzeichnet.

Die Einbettung in die „natürliche“ sprachliche Umwelt und die schulische Ausbildung scheinen in beiden Bereichen für etwa 2/3 der deutschen Schülerinnen und Schüler bis zum Ende der Klasse neun solide Grundfähigkeiten zu vermitteln.

Sorge bereiten dagegen die 23% der Schreibenden, die für die Pragmatik und die 35%, die für die Sprachsystematik auf Niveau A verortet sind und damit zwar einen Brief mit wenigen Informationen zum Thema verfassen können, aber in einer solch niedrigen sprachlichen Qualität, dass zu befürchten steht, dass das kommunikative Ziel nicht erreicht wird. Noch größere Probleme haben die 6% der Schreibenden für die semantisch-pragmatische und 3% für die systematische Dimension, die sich unter dem Niveau A befinden. Diese Schreibenden produzieren entweder gar keine Texte oder Texte in einer Qualität, die noch nicht mit ausreichender Sicherheit eine

Zuordnung zum ersten Niveau erlaubt, also sehr kurze, wenig aussagekräftige Texte ohne unterscheidbare Textelemente. Diese sind ohne bzw. mit minimalem Bezug zu Informationen der Aufgabenstellung, stilistisch unangemessen in sehr expressivem Wortschatz. Oder es werden sinnentstellende Texte mit elementaren Fehlern in der Wortgrammatik und einem vorrangig aus Reihungen bestehenden Satzbau, meist ohne Interpunktion produziert. Diese Schreiben können keine (positive) Reaktion des Empfängers bewirken und verfehlen damit ganz sicher ihr kommunikatives Ziel. Hier müssen für den schulischen Unterricht weitere Fördermaßnahmen konzipiert werden.

Schülerleistungen – bildungsgangspezifisch

Betrachtet man die Ergebnisse differenziell nach den einzelnen Bildungsgängen, so zeigt sich, konsistent über beide Teildimensionen, folgender Befund: Die Gymnasiaschüler schreiben im Mittel überdurchschnittliche Texte (Mittelwert 576 Punkte für Semantik/Pragmatik und 569 für Sprachsystematik), die Realschüler mit jeweils 511 Punkten in beiden Teildimensionen etwa durchschnittliche. Deutlich schwächere Texte (455 bzw. 446 Punkte) wurden an den IGS und von den Hauptschülern (432 bzw. 428 Punkte) produziert. Anhand der Verteilungen auf die Kompetenzniveaus für die einzelnen Bildungsgänge, wie sie Abbildung 8.3 zeigt, lassen sich zwei Profilgruppen identifizieren.

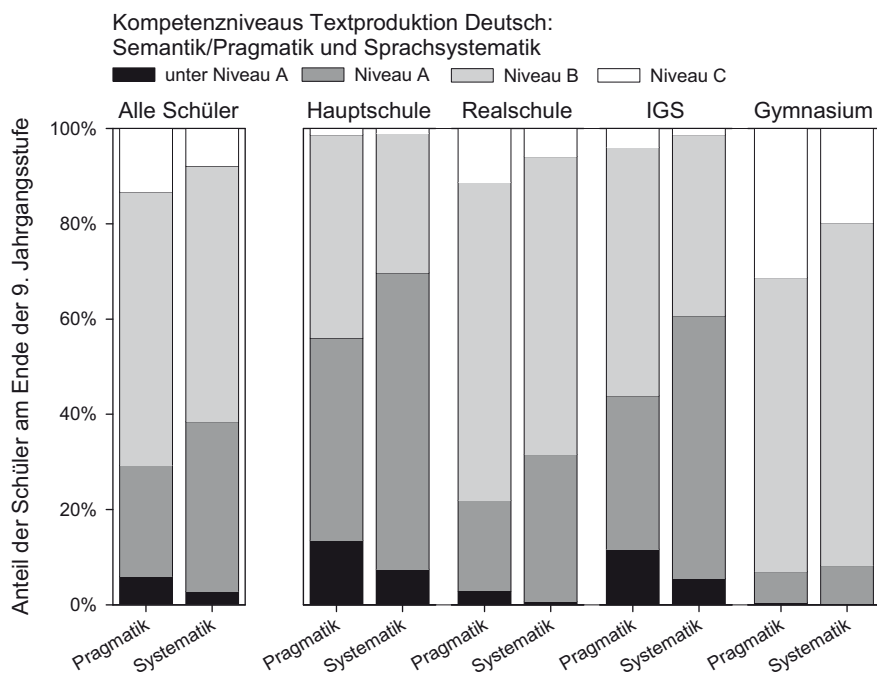


Abbildung 8.3: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Textproduktion Deutsch: Semantik/Pragmatik“ und „Sprachsystematik“ am Ende der neunten Jahrgangsstufe unter Beachtung der Bildungsgänge.

Lediglich an den Realschulen (0.1% Semantik/Pragmatik und 0.2% Systematik) und den Gymnasien (0.4% vs. 0.7%) gelingt es überhaupt, „meisterliche“ Briefe zu verfassen. Darüber hinaus übersteigt der Prozentsatz der Schülerinnen und Schüler auf und oberhalb des jeweiligen Kompetenzniveaus B an diesen Schulen die Zwei-Drittel-Marke (RS: 78%/69%; GYM: 93%/92%), wobei an den Gymnasien die Anteile deutlich höher sind. Andererseits muss festgestellt werden, dass immerhin noch 22% der Realschüler und 7% der Gymnasiasten im semantisch-pragmatischen und 31% der Realschüler und 8% der Gymnasiasten im sprachsystematischen Bereich das Kompetenzniveau A nicht überschritten haben, womit eine zielgerichtete erfolgreiche Kommunikation nur noch eingeschränkt gesichert werden kann.

Die zweite Gruppe stellen die IGS und Hauptschulen dar, in denen lediglich 56% bzw. 44% in der semantisch-pragmatischen Komponente und sogar nur 40% bzw. 31% in der sprachsystematischen Komponente inner- bzw. oberhalb des Kompetenzniveaus B zu finden sind. Der Anteil der Schülerinnen und Schüler, die Texte auf Kompetenzniveau A (Semantik/Pragmatik: 32% vs. 42%; Systematik: 54% vs. 62%) und auch darunter (12% vs. 14%; 6% vs. 7%) produzieren, die also im Bereich schriftlicher Kommunikation nur mit Einschränkungen oder gar nicht mehr erfolgreich sein können, macht deutlich, dass hier große Defizite vorzufinden sind.

Über alle Bildungsgänge hinweg zeichnen sich in der sprachsystematischen Komponente größere Defizite ab als in der semantisch-pragmatischen. Die Leistungen der Neuntklässler sind im letzteren Bereich hinsichtlich des kommunikativen Erfolges offenbar besser entwickelt, lediglich für die Anzahl der unter Niveau A liegenden Texte gilt dies nicht. Hier scheitert die zielgerichtete Kommunikation offensichtlich eher an einer falschen „Grundhaltung“ und der darin geschuldeten übermäßigen Expressivität oder einer falschen Empfängerorientierung als an einer nicht mehr erschließbaren sprachlichen Umsetzung.

Kompetenzzuwachs in der neunten Jahrgangsstufe

Insgesamt findet ein Kompetenzzuwachs lediglich in der semantisch-pragmatischen Komponente, und zwar um etwa fünf Punkte, statt, nicht aber in der sprachsystematischen Komponente (ohne Abbildung). Die Veränderungen innerhalb der neunten Klasse lassen sich dabei differenzierter für die einzelnen Bildungsgänge zeigen, wie aus Abbildung 8.4 ersichtlich wird.

Insgesamt findet in der semantisch-pragmatischen Teilkompetenz ein Leistungszuwachs statt, der größenordnungsmäßig im unteren Bereich der Entwicklungen in anderen längsschnittlich untersuchten Kompetenzen in Deutsch liegt. Dass er im Realschulbildungsgang mit 8.4 Punkten etwas stärker ausfällt als im gymnasialen (3.8 Punkte) und Hauptschulbildungsgang (3.2) ist insofern erwartungskonform, als dort ein Unterricht zu effektivem Schreiben im Format eines Briefes verbreiteter sein dürfte als in den anderen Bildungsgängen.

In allen Bildungsgängen, auch an den oben nicht genannten IGS, finden Verbesserungen vor allem im unteren Leistungsbereich statt, wobei sich der prozen-

tuale Anteil der auf bzw. unter Niveau A Schreibenden in den Realschulen und den Gymnasien, in denen er ohnehin geringer ist, mit -5% bzw. -2% deutlicher verringert als an den Hauptschulen (-0.6%) und den IGS (-0.2%). Der Leistungszuwachs der Hauptschulen beruht auf einer deutlicheren Abnahme der Schreibenden unter Niveau A (-4%).

Es zeigt sich hier ansatzweise, dass der Schwerpunkt der Schreibausbildung in der neunten Jahrgangsstufe vorrangig im semantisch-pragmatischen Bereich liegt. Für die Realschülerinnen und -schüler zielt er offensichtlich auf eine bessere Zielorientierung von Texten, die vom Empfänger komplikationslos zu entschlüsseln sein müssen, während für die Hauptschülerinnen und -schüler wenigstens zielgerichtete, entschlüsselbare Schreiben angestrebt werden. Im oberen Leistungsbereich dagegen scheint in der neunten Jahrgangsstufe eine institutionelle Schreibförderung schwieriger umzusetzen zu sein, hier lassen sich keine Leistungszuwächse nachweisen, genauso wenig wie für den gesamten Bereich der Sprachsystematik. Minimale Veränderungen in diesem Bereich können auch auf größere Schreibroutine durch den Schul- und Lebensalltag zurückzuführen sein.

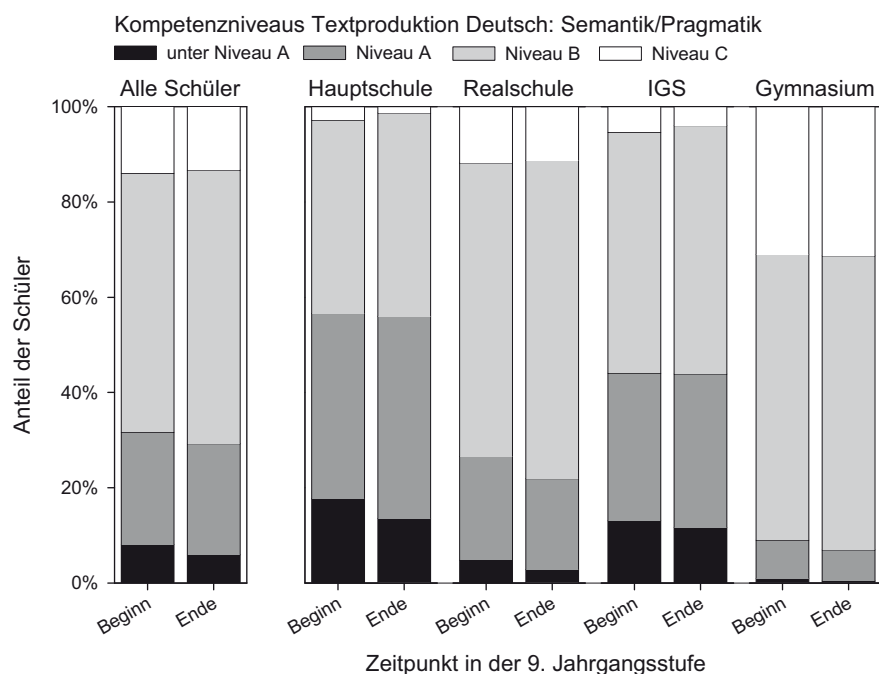


Abbildung 8.4: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Textproduktion Deutsch: Semantik/Pragmatik“ zu Beginn und am Ende der neunten Jahrgangsstufe unter Beachtung des Bildungsgangs.

Geschlechtsspezifische Ergebnisse

In den Auswertungen der DESI-Textproduktionsaufgaben bestätigt sich der allgemein bekannte Befund der Überlegenheit der Mädchen gegenüber den Jungen in schriftsprachlichen Tätigkeiten, wobei der Abstand in der semantisch-pragmatischen Komponente größer ist (Differenz von 74 Punkten) als in der sprachsystematischen (Differenz von 49 Punkten). In beiden Fällen bestehen dabei die größten Divergenzen in den Realschulen, in denen Mädchen verhältnismäßig stärker als in anderen Bildungsgängen vom Schreibunterricht profitieren als die Jungen. Der allgemeine Leistungszuwachs in der semantisch-pragmatischen Komponente geht vor allem auf die Mädchen zurück, die innerhalb der neunten Jahrgangsstufe um durchschnittlich 11 Punkte besser werden, während die Leistungen der Jungen durchschnittlich stagnieren, an den IGS mit -9 Punkten und an den Gymnasien mit -4 Punkten sogar schlechter werden. Für die sprachsystematische Komponente ist dagegen weder für die Jungen noch für die Mädchen eine Leistungsveränderung zu verzeichnen. Für beide Teilkompetenzen lohnt dabei ein differenzierter Blick auf die Gruppen mit unterschiedlichen sprachlichen Voraussetzungen.

Effekte der Erstsprache

Deutschsprachige Schülerinnen und Schüler sind durchschnittlich in beiden Teilkompetenzen am erfolgreichsten (Pragmatik: 512/Systematik: 512), gefolgt von den mehrsprachig Aufwachsenden (495/492) und den Lernenden mit nicht deutscher Erstsprache (443/445). Die Unterschiede zwischen den so gebildeten Gruppen sind die größten auffindbaren im DESI-Modul Textproduktion überhaupt. Differenzielle Effekte in den einzelnen Bildungsgängen lassen sich hingegen für beide Teilkompetenzen statistisch signifikant nur noch in der Unterscheidung zwischen deutschsprachigen vs. nicht deutschsprachigen Schreibenden nachweisen. Eine Ausnahme bilden hier die Gymnasien, an denen auch für die mehrsprachig aufwachsenden Jugendlichen statistisch bedeutsame Unterschiede auffindbar sind. Insgesamt lässt sich zeigen, dass in beiden Teilkompetenzen die nicht deutschsprachigen Gymnasiasten deutlich besser (Pragmatik: 533 Punkte/Systematik: 552) sind als die beste Gruppe der deutschsprachigen Realschülerinnen und Realschüler (515/515). Ob sich darin ein anspruchsvollerer Unterricht oder aber ein vorher stattgefundener Segregationseffekt aufgrund hoher sprachlicher Eingangsvoraussetzungen für das Gymnasium widerspiegelt, kann hier nicht entschieden werden.

8.7 Implikationen für den Deutschunterricht

Die Befunde zur Textproduktion Deutsch zeigen, dass

- insgesamt etwa von zwei Dritteln der Schreibenden eine textpragmatische und sprachsystematische Kompetenz erreicht wird, die eine zielgerichtete schriftliche Kommunikation ermöglicht;

- der Prozentsatz schwer oder gar nicht zu entschlüsselnder Texte dagegen an den Hauptschulen und den IGS (mit etwa 50%) sehr hoch ist;
- Leistungszuwächse innerhalb der neunten Jahrgangsstufe nur im semantisch-pragmatischen Bereich zu verzeichnen sind und dort auch ausschließlich auf Fördermaßnahmen im unteren Leistungsbereich zurückzuführen sind und
- lediglich Mädchen und mehrsprachig aufwachsende Schülerinnen und Schüler signifikant besser werden, wobei eine hohe Stabilität der bildungsgangsspezifischen Muster zu erkennen ist.

Angesichts dieser Ergebnisse erscheinen Rufe nach deutlich stärkerer Förderung der Schreibfähigkeiten der Schülerinnen und Schüler als hoch plausibel. Der Deutschunterricht darf sich offenkundig nicht mit einem Verharren auf einem „irgendwie verständlichen“ Text begnügen, sondern muss neben der Reduktion des Anteils der Schülerinnen und Schüler auf oder unterhalb von Niveau A auch Verbesserungen im oberen Leistungsbereich anstreben.

Dass innerhalb einer Jahrgangsstufe in der sprachsystematischen Komponente von Texten keine Leistungszuwächse zu verzeichnen sind, sollte bei einem hohen Anteil an unverständlichen oder nur eingeschränkt verständlichen Texten alarmieren. Das Herauslösen oder Heruntergewichten von Anforderungen an die orthographischen bzw. grammatischen Fähigkeiten der Schreibenden bei der Bewertung von Abschlussklausuren, wie es derzeit in vielen Bundesländern diskutiert wird, erscheint angesichts solcher Ergebnisse eher kontraproduktiv.

Literatur

- Brinker, K. (1988): Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin: Schmidt Verlag.
- deBeaugrande, R./Dressler, W. (1981): Einführung in die Textlinguistik. Tübingen: Niemeyer.
- Harsch, C./Lehmann, R. H./Neumann, A./Schröder, K. (2007): Schreibfähigkeit. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band I. Weinheim: Beltz, S. 42-62.
- Hartig, J. (2007): Skalierung und Kompetenzniveaus. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 83-99.
- Lehmann, R. H. (1988): Reliabilität und Generalisierbarkeit der Aufsatzbeurteilung im Rahmen des Hamburger Beitrags zur internationalen Aufsatzstudie der IEA. In: Empirische Pädagogik. Zeitschrift zu Theorie und Praxis erziehungswissenschaftlicher Forschung 2/4, S. 349-365.
- Lehmann, R. H. (1990): Aufsatzbeurteilung – Forschungsstand und empirische Daten. In: Ingenkamp, K./Jäger, R. (Hrsg.): Tests und Trends. Jahrbuch der pädagogischen Diagnostik, Bd. 8. Weinheim: Beltz, S. 64-94.
- Linke, A./Nussbaumer, M./Portmann, P. R. (1996): Studienbuch Linguistik. Tübingen: Niemeyer.
- Neumann, A. (2007a): Testverweigerung in DESI-Textproduktion: Nutzen der offenen Antwortformate zum Ausdruck der individuellen (Nicht-)Interessen. In: Gailberger, S./Krelle, M. (Hrsg.): „Wissen und Kompetenz“. Entwicklungslinien und Kontinuitäten in der Deutschdidaktik. Festschrift anlässlich des 65. Geburtstages von Heiner Willenborg. Hohengehren: Schneider-Verlag, S. 185-193.

- Neumann, A. (2007b): Briefe schreiben in Klasse 9 und 11. Beurteilungskriterien, Textstrukturen und Schülerleistungen. Urspr. Titel der Dissertation an der Universität Hamburg: Dimensionen der Schreibfähigkeit. Differenzierte Analysen der Texte aus DESI und LAU11/ULME1. Münster/New York/München/Berlin: Waxmann Verlag.
- Neumann, A. (2006): Stabilität von Raterurteilen über die Zeit – Anpassung an vorhandene Schülerleistungen. Auswertung von Replikationsstudien zu den Urteilen in DESI-Textproduktion. In: Empirische Pädagogik 20, H. 3, S. 286-296.
- Nussbaumer, M. (1991): Was Texte sind und wie sie sein sollen. Tübingen: Niemeyer.
- Wirtz, M./Caspar, F. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen/Bern/Toronto/Seattle: Hogrefe.

Günther Thomé / Wolfgang Eichler

9 Rechtschreiben Deutsch

9.1 Testkonstrukt

Eine rein quantitative Auswertung orthographischer Leistungen kann unter lernpsychologischen und didaktischen Aspekten heute nicht mehr als zeitgemäß angesehen werden. Orthographische Fehler in schriftlichen Äußerungen sind fast ausschließlich als Ausdruck von noch nicht abgeschlossenen Lernprozessen zu sehen. Dieses gilt unabhängig von Lebensalter und Bildungsbiographie, denn auch nach mehrjährigem schulischem Unterricht ist es durchaus möglich, dass schriftsprachliche Lernprozesse teilweise ausgeblieben oder missglückt sind.

Die orthographischen Lernprozesse, die während der Schulzeit ablaufen, können mittlerweile durch qualitative Fehleranalysen recht gut rekonstruiert werden (vgl. Eichler 1983; May 2000; Herné/Naumann 2001; Thomé/Thomé 2004). Hier soll explizit auf die Oldenburger Fehleranalyse Bezug genommen werden, deren Fehlertypologie aus Tabelle 9.1 hervorgeht.

So gelten bestimmte Fehlertypen, die einen nennenswerten Anteil an der Gesamtfehlerzahl ausmachen, als Leitphänomene der einzelnen Entwicklungsstufen im Schrifterwerb (Thomé 2003). Diese Fehlertypen können überindividuell definiert und in Bezug auf den Lerngegenstand, die Orthographie, systematisch klassifiziert werden.

Tabelle 9.1: Fehlertypen nach der Oldenburger Fehleranalyse (OLFA).

OLFA - Oldenburger Fehleranalyse- Version 2.1 © G.Thomé/D.Thomé 2004		Gruppe I Vor- bis Protoalph.	Gruppe II Alphabet.	Gruppe III Orthograph.
01	Klein- für Großschreibung			
02	Groß- für Kleinschreibung			
03	Großschreibung im Wort			
04	Getrennt- für Zusammenschreibung			
05	Zusammen- für Getrenntschreibung			
06	Getrenntschreibung von unselbstständigen Teilen			
07	Einfachschreibung für Verdopplung			
08	Verdopplung für Einfachschreibung bei Kurzvokal			
09	Einfachschreibung für markierte Länge			
10	Mark. Längenschr. für Einfachschr. bei Langvokal			
11	Verdopplung nach Langvokal oder Konsonant			
12	Markierte Länge bei Kurzvokal			
13	s für ß			
14	ß für s			
15	ss für ß			
16	ß für ss			
17	e/eu für ä/äu			
18	ä/äu für e/eu			
19	p t k für b d g im Silbenendrand			
20	b d g für p t k im Silbenendrand			
23	f für v			
24	v für f			
25	w für v			
26	v für w			
27	ch für g im Silbenendrand			
28	g für ch im Silbenendrand			
29	Konsonantenzeichen fehlt			
30	Konsonantenzeichen zugefügt			
31	Vokalzeichen fehlt			
32	Vokalzeichen zugefügt			
33	Falscher Konsonant			
34	Falscher Vokal			
35	Zeichenumstellung			
36	Umlautbezeichnung			
37	Sonstige Fehler (auch Fremdwortfehler)			
Fehlertypen absolut nach Hauptgruppen				
Gesamtfehlertypen:		Fehler in %		
Anzahl der Wörter:		Fehler auf 100 Wörter:	KWA:	KWB:

Bei der Klassifikation von Rechtschreibfehlern kann man sowohl unter lernpsychologischem als auch unter sachstrukturellem Aspekt vier Hauptkategorien unterscheiden:

1: Wortschreibungen, denen nicht die korrekte Lautstruktur des betreffenden Wortes zu entnehmen ist (in der OLFA-Liste die Nummern 11, 12, 15, 16 und 29-35).

Diese Fehler kommen zustande durch Auslassen oder Hinzufügen von Graphemen (z. B. *Geburstag für Geburtstag, *artmen für atmen) und durch die Schreibung eines Konsonant- oder Vokalgraphems, das im deutschen Orthographiesystem nicht mit dem zu schreibenden Phonem korrespondiert (z. B. *wrist für frißt, *Strößchen für Sträußchen). Diese Fehler können grundsätzlich an jeder Stelle eines Wortes auftreten. Das betroffene didaktische Feld bei vielen Fehlern dieser Art ist die Lautebene.

2: Wortschreibungen, die zwar die korrekte Lautstruktur des betreffenden Wortes wiedergeben, aber nicht die geforderte orthographisch markierte Schreibung enthalten (in der OLFA-Liste die Nummern 07, 09, 13, 17, 19, 23, 25, 27).

Beispiele für diese Fehlertypen sind *fält für fällt, *bükt für bückt (Fehlertyp 07), *faren für fahren (Fehlertyp 09), *Ferkeuferin für Verkäuferin (Fehlertypen 23 und 17), *Walt für Wald (Fehlertyp 19). Diese Fehler können nur an Stellen auftreten, an denen ein Orthographem zu schreiben ist, was auf etwas mehr als 5% aller Grapheme zutrifft. Das didaktische Feld bei Fehlern dieser Art bilden Phonologie, Morphologie und Lexik.

3: Fehler in der Großschreibung (in der OLFA-Liste die Nummern 01 und 02; Fehlertyp 03 nach dem dritten Schuljahr nicht mehr nachgewiesen).

Zu dieser Kategorie gehören die Kleinschreibung von groß zu schreibenden Wörtern (Nomen und nominalisierte Formen, Eigennamen, Anredepronomen in der Höflichkeitsform und Satzanfänge) sowie die Großschreibung von klein zu schreibenden Wörtern. Das betroffene didaktische Feld bilden Lexik und Syntax.

4: Fehler in der Zeichensetzung (in der OLFA-Liste nicht enthalten).

In diese Kategorie fällt besonders die Kommasetzung, aber auch fehlende Punkte am Satzende, Anführungszeichen bei direkter Rede sowie seltenere Phänomene wie Semikolon und Doppelpunkt können Fehler verursachen. Das didaktische Feld bilden hier die syntaktische und die Textebene.

9.2 Testaufgaben

Im DESI-Test zur Rechtschreibung im Deutschen wurde die orthographische Kompetenz durch ein eigens konstruiertes Diktat erfasst. In der zur Verfügung stehenden Testzeit von zehn Minuten sollten auch langsam schreibende Schüler den gesamten Text schreiben und noch einmal zur Korrektur durchsehen können. In Vortests zeigte sich, dass ein Text mit 68 Wörtern diesen Anforderungen entspricht. Trotz der Kürze des Diktattextes sind in diesem alle Rechtschreibphänomene vertreten und damit getestet worden. So sind jeweils mehrere Wörter enthalten, in denen sich die orthographischen Bereiche Groß- und Kleinschreibung, Getrennt- und Zusammenschreibung,

Konsonantenverdoppelung, das/dass, Dehnungs-h, Umlaute und Auslautverhärtung, seltene Grapheme wie <v>, <dt> u. Ä. sowie Konsonantenhäufungen finden und drei Fremdwörter, die im Bekanntheits- und Schwierigkeitsgrad etwa mit „Garage“ vergleichbar sind.

In der Hauptuntersuchung musste kein Schüler wegen des Tempos das Schreiben abbrechen. Der Text wurde von Tonbandkopien vorgespielt, um Einflüsse durch unterschiedliches Diktieren auszuschließen.

Mit der Verwendung eines Diktates in DESI ist keine positive Bewertung des Diktates als Unterrichtsinstrument verbunden. Während durch die große Zahl der untersuchten Schülerinnen und Schüler diese Testform notwendig war, sollten Diktate in der Schule möglichst selten geschrieben werden. Da die Rechtschreibleistung in DESI nur zum zweiten Testzeitpunkt erhoben wurde, können keine Aussagen über eventuelle Entwicklungen innerhalb eines Jahres gemacht werden.

9.3 Kompetenzniveaus

Alle im DESI-Diktat festgestellten Rechtschreibfehler wurden qualitativ erfasst und einer der o. g. Kategorien zugeordnet. Auf dieser Grundlage konnten mit Hilfe der Rasch-Skalierung (vgl. Kapitel 3) für die genannten Orthographiebereiche Schwierigkeitsparameter berechnet und Schwellen für die einzelnen Niveaus bestimmt werden.

Die Zuordnung einer Schülerin/eines Schülers zu einem Kompetenzniveau geschieht unter der Bedingung, dass mindestens 65% der für dieses Niveau typischen Anforderungen bewältigt werden. Die Beschreibung der Kompetenzniveaus A bis D ist in Tabelle 9.2 zusammengestellt.

Tabelle 9.2: Beschreibung der Kompetenzniveaus in „Rechtschreiben Deutsch“ mit Fehlerbeispielen.

Kompetenz-niveau	Beschreibung	Itembeispiele
D (mehr als 641 Punkte)	Zusätzlich wird die Zeichensetzung weitgehend korrekt geleistet.	In diese Kategorie fällt besonders die Kommasetzung, aber auch Punkte am Satzende oder Anführungszeichen bei direkter Rede sowie seltenere Phänomene wie Semikolon und Doppelpunkt.
C (582-640 Punkte)	Zusätzlich werden Anforderungen im Bereich der Groß- und Kleinschreibung weitgehend gemeistert.	Zu dieser Kategorie gehören die Kleinschreibung von groß zu schreibenden Wörtern (Nomen und nominalisierte Formen, Eigennamen, Anredepronomen in der Höflichkeitsform und Satzanfänge) sowie die Großschreibung von klein zu schreibenden Wörtern.
B (382-581 Punkte)	Zusätzlich sind die Wortschreibungen in den Hauptbereichen, wie Konsonantenverdoppelung, Dehnungs-h, Morphemkonstanz bei Auslautverhärtung und Umlautung usw. weitgehend korrekt.	Als Fehler treten kaum noch oder nicht mehr auf: z.B. *fällt für fällt, *bükt für bückt, *faren für fahren, *Ferkueferin für Verkäuferin, *Walt für Wald.
A (344-381 Punkte)	Die grundlegenden Phonem-Graphembeziehungen werden weitgehend korrekt umgesetzt.	Als Fehler treten kaum noch oder nicht mehr auf: Auslassen oder Hinzufügen von Graphemen z.B. *Geburstag für Geburtstag, *artmen für atmen; Schreibung eines Konsonant- oder Vokalgraphems, das im deutschen Orthographiesystem nicht mit dem zu schreibenden Phonem korrespondiert z. B. *wrist für frißt, *Strößchen für Sträußchen.

Schülerinnen und Schüler, die Kompetenzniveau A nicht erreichen, haben nicht nur Probleme in allen orthographischen Hauptbereichen, sondern produzieren auch Fehler in der grundlegenden Phonem-Graphem-Zuordnung. Der normierte Leistungswert liegt unter 344.

9.4 Ergebnisse

Leistungsverteilung am Ende der neunten Jahrgangsstufe

Die Abbildung 9.1 zeigt die Verteilung der Kompetenzniveaus für alle Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe. In Tabelle 9.3 sind zusätzlich die absoluten Fehlerzahlen angegeben, die von Schülerinnen und Schülern auf den einzelnen Kompetenzniveaus gemacht wurden.

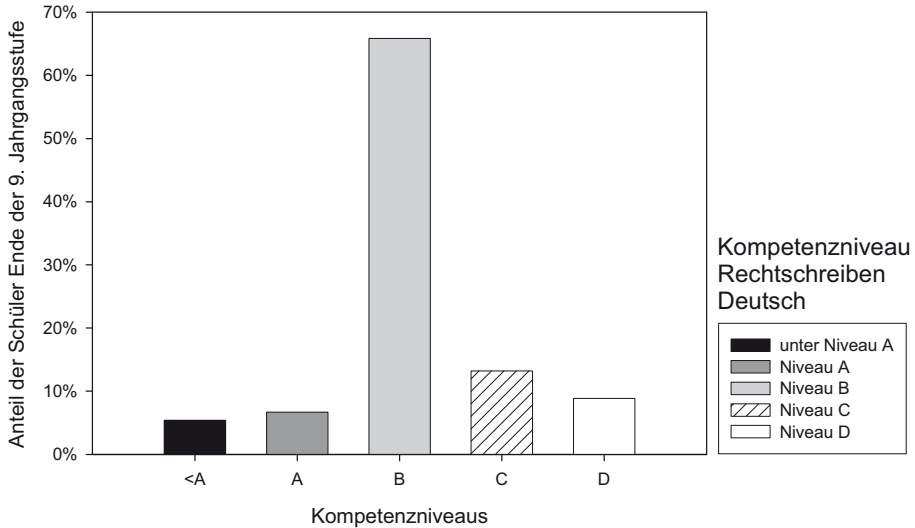


Abbildung 9.1: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Rechtschreiben Deutsch“ am Ende der neunten Jahrgangsstufe.

Insgesamt sind die Werte symmetrisch auf die Niveaus verteilt; die mittlere Gruppe mit Kompetenzniveau B ist sehr groß. Die Schreibungen dieser Schülerinnen und Schüler sind in den Hauptbereichen korrekt. Der Anteil der Schüler mit einem geringeren Kompetenzniveau (A und <A) beträgt zusammen 12% (vgl. auch Tabelle 9.3). Diese Schüler haben erhebliche Probleme in der Rechtschreibung. Lediglich die Schüler mit höherem Kompetenzniveau (C und D) können als kompetente Rechtschreiber bezeichnet werden. Ihr Anteil an der Gesamtzahl der Schüler beträgt nur 22%.

Die in Tabelle 9.3 dargestellten absoluten Fehlerwerte lassen unter quantitativer Perspektive die insgesamt schwachen Leistungen im Rechtschreiben erkennen. Auch wenn der Diktattext als orthographisch eher schwer eingestuft werden kann, dürften im schulischen und beruflichen Kontext bei mittlerem Leistungsanspruch die orthographischen Fähigkeiten von etwa der Hälfte der Schüler als nicht ausreichend bewertet werden.

Tabelle 9.3: Verhältnis von orthographischem Kompetenzniveau und absoluter Fehlerzahl (Graphemfehler plus Interpunktionsfehler) aus dem DESI-Diktat mit 68 Wörtern.

Kompetenz-niveau	Mittelwert der absoluten Fehlerzahl	Median	Anteil der Schüler
< A	41.5	39	5.4%
A	32.6	30	6.7%
B	16.4	14	65.9%
C	7.5	7	13.2%
D	5.1	4	8.9%

Orthographische Kompetenzen nach Bildungsgang

Die Abbildung 9.2 zeigt die Verteilung der orthographischen Kompetenzniveaus auf die Bildungsgänge. Mit über 58% in Kompetenzniveau C oder D kann für das Gymnasium eine relativ gute Leistung festgestellt werden. Bezieht man den Mittelwert von 16.4 Fehlern (auf 68 Wörter) für das Kompetenzniveau B mit ein, wird deutlich, dass etwa die Hälfte der Realschüler und ein großer Teil der Haupt- und Gesamtschüler deutliche Defizite in der Rechtschreibung haben. Kompensatorische Maßnahmen sind daher in diesen Fällen unbedingt erforderlich.

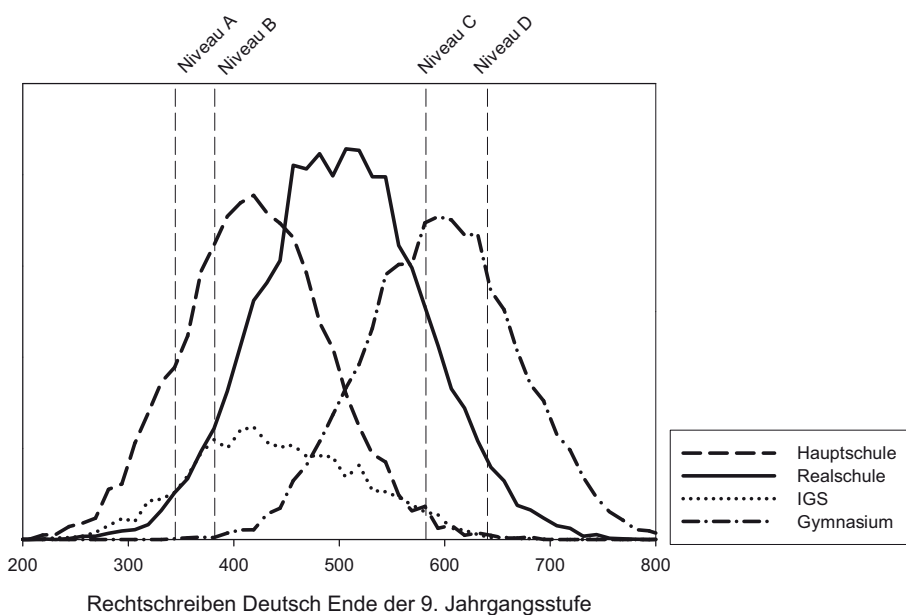


Abbildung 9.2: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Rechtschreiben Deutsch“ am Ende der neunten Jahrgangsstufe unter Beachtung der Bildungsgänge.

9.5 Implikationen für den Deutschunterricht

Die Ergebnisse machen deutlich, dass das Thema Rechtschreibung in der neunten Jahrgangsstufe keineswegs erledigt ist. Dieses deckt sich auch mit Beobachtungen in den weiterführenden berufsbildenden Schulen, wo für bestimmte Klassen der Rechtschreibunterricht gegenwärtig regelrecht „wiederentdeckt“ wird.

Die in der Schule oft geübte Praxis, bei Aufsätzen die Rechtschreibfehler nicht oder nur fakultativ nebenbei zu bewerten, kommt hier zu einem wenig rühmlichen Ende. Gute bis mittlere Leistungen in der Rechtschreibung sind für Arbeitgeber eine wichtige Voraussetzung für die Einstellung oder ein Ausbildungsverhältnis.

Literatur

- Eichler, W. (1983): Kreative Schreibirrtümer. Zur Auseinandersetzung des Schülers mit dem Verhältnis Laut-Schrift und mit den Rechtschreibregeln. In: Diskussion Deutsch 14, H. 74, S. 629-640.
- Herné, K. L./Naumann, C. L. (⁵2001): Aachener Förderdiagnostische Rechtschreibfehler-Analyse (AFRA). Aachen: Alfa Zentaurus.
- May, P. (⁵2000): Diagnose orthographischer Kompetenz. Zur Erfassung der grundlegenden Rechtschreibstrategien mit der Hamburger Schreib-Probe. Unter Mitarbeit von U. Vieluf, V. Malitzky. Handbuch. Hamburg: Verlag für pädagogische Medien.
- Thomé, G. (2003): Entwicklung der basalen Rechtschreibkenntnisse. In: Bredel, U./Günther, H./Klotz, P./Ossner, J./Siebert-Ott, G. (Hrsg.): Didaktik der deutschen Sprache, Band 1. Paderborn: Schöningh, S. 369-379.
- Thomé, G./Thomé, D. (2004): OLFA – Oldenburger Fehleranalyse: Instrument und Handbuch zur Ermittlung der orthographischen Kompetenz aus freien Texten ab Klasse 3 und zur Qualitätssicherung von Fördermaßnahmen. Version 2.0. Oldenburg: Igel Verlag Wissenschaft (zu beziehen über: www.isb-oldenburg.de).

Wolfgang Eichler

10 Sprachbewusstheit Deutsch

10.1 Testkonstrukt

Im DESI-Modul Sprachbewusstheit Deutsch geht es darum festzustellen, inwieweit die Jugendlichen in der Lage sind, grammatisch exakt und stilbewusst mit Sprache in für dieses Schulalter wesentlichen Bereichen umzugehen. Der Test prüft sowohl die explizite Kenntnis grammatischer Kategorien, die in der Fachdidaktik als „deklaratorisches Wissen“ bezeichnet wird, als auch das so genannte „monitoring“, d.h. die laufende Kontrolle und ggf. Korrektur des eigenen und fremden Sprachgebrauchs hinsichtlich grammatischer und stilistischer Angemessenheit. Dieses „monitoring“ ist bei einfachen Fehlerkontrollen vermutlich stark automatisiert. Mit diesem Konzept von Sprachbewusstheit, bei dem prozedurale Komponenten im Vordergrund stehen und nicht nur die explizite Kenntnis von Regeln untersucht wird, knüpft DESI an die neuere *language awareness*-Forschung an (vgl. u.a. Andresen/Funk 2003; Eichler 2004a, b; Eichler/Nold 2007; Luchtenberg 1995).

Die theoretische Begründung der Kompetenzniveaus erfolgt in Anlehnung an das Modell von Karmiloff-Smith (1992). Da das Kompetenzkonzept Sprachbewusstheit bereits an anderer Stelle sehr ausführlich nachzulesen ist, wird an dieser Stelle auf theoretische Erläuterungen zugunsten der Ergebnisdarstellung verzichtet (vgl. Eichler 2007; Eichler/Nold 2007).

10.2 Testaufgaben

Sprachbewusstheit in DESI erfasst unterschiedliche, in verschiedenen Sprachgebrauchssituationen immer wieder benötigte Teilkompetenzen. Die prozedurale und explizite Sprachbewusstheit spielt auch in anderen Modulen von DESI eine gewichtige Rolle, z.B. phonologische und orthographische Bewusstheit in der Rechtschreibung, Textsorten-, Adressaten-, und Stilbewusstheit im Modul Textproduktion, und in vielen Bereichen des Fremdsprachenerwerbs (vgl. dazu Eichler/Nold 2007). Konkret sprechen die Testaufgaben drei Arten von sprachlichen Phänomenen an:

- schwierige, spät und bewusst gelernte sowie durch den Sprachwandel gefährdete grammatische Phänomene (z.B.: Unterscheidung Dativ-/ Akkusativobjekt, Kongruenzen innerhalb eines Satzes – d.h. Übereinstimmungen hinsichtlich Kasus, Genus, Numerus, Person – über einen größeren Fokusbereich hinweg, Gebrauch des Genitivobjekts, Konjunktivformen),
- Aspekte der Stilistik (die Einbettung fester Wortverbindungen – so genannte Kollokationen –, Stilverträglichkeit, Stilistik des Konjunktivs 1 und 2) und
- Wissensbestände im Bereich der indirekten Rede und des Konjunktivs.

Die Gegenstandsbereiche stimmen mit ausgewählten Inhalten der Rahmenrichtlinien im Bereich „Sprachreflexion“ und auch mit einer Reihe von Vorgaben der Bildungsstandards im Bereich „Sprache und Sprachgebrauch untersuchen“ für den Mittleren Schulabschluss überein, und sie spiegeln typische Fehler von Jugendlichen der neunten Jahrgangsstufe wider (vgl. u.a. Eichler 2003 und Eichler/Nold 2007). Je nach Aufgabenstellung sind Fehler zu korrigieren, grammatische Begriffe zuzuordnen oder kleine kreative Formulierungen zu leisten.

10.3 Kompetenzniveaus

Die Beschreibung der Kompetenzniveaus A bis E sowie charakteristische Beispielaufgaben für jedes Kompetenzniveau sind in Tabelle 10.1 zusammengestellt. Jedes höhere Niveau schließt die Fähigkeiten der darunter liegenden Niveaus mit ein.

Tabelle 10.1: Beschreibung der Kompetenzniveaus in „Sprachbewusstheit Deutsch“.

Kompetenz-niveau	Beschreibung	Beispiele
E (ab 709 Punkten)	Aktive Anwendung deklaratorischen Wissens	Kreuze an: Im folgenden Satz ist die Handlung in ihrer sprachlichen Form als wirklich, möglich oder unwirklich dargestellt. „Wenn ich die 100 m in 11,2 Sekunden gelaufen wäre, wäre ich Jahresbester geworden.“ <input type="checkbox"/> wirklich <input type="checkbox"/> möglich <input type="checkbox"/> unwirklich
D (613 bis 708 Punkte)	Entfaltetes gramma- tisches und stilistisches monitoring: Komplexe Stilfehler erkennen und beseitigen, Mehrdeutigkeiten auflösen.	Wer soll hier überquert werden? Verbessere den folgenden Satz so, dass er eindeutig ist. <i>Zebrastreifen sollen das Überqueren der Passanten sichern.</i>
C (550 bis 612 Punkte)	Entfaltetes gramma- tisches monitoring: Sicherheit im Erkennen und Korrigieren auch schwieriger grammatischer Phänomene. Einfache Zuordnung grammatischer Begriffe zu Phänomenen.	Im folgenden Satz ist eine Form grammatisch falsch gebildet. Unterstreiche das Wort (die Wörter) mit der falschen Form. Versuche dann eine Verbesserung. <i>Viele Jugendliche schämen sich ihrem Aussehen.</i> Unterstreiche die Formen des Konjunktivs im folgenden Satz: „Wenn ich nicht noch zur Bank gelaufen wäre, wäre ich nicht zu spät gekommen.“
B (488 bis 549 Punkte)	Einfache gramma- tische und stilistische Sprachbewusstheit: Stil und Kohärenz in einfachen Kontext- und Inhaltsbereichen herstellen.	Schreibe mit den folgenden Stichworten eine kleine Zeitungsnotiz für eine Tageszeitung wie die Northwest Zeitung, Die Welt, die Süddeutsche Zeitung. <i>Großes Publikum – von 1995 – Die Fantastischen Vier – Erfolg – Live – Hip-Hop – Auftritt in der Stadthalle – Band</i>
A (425 bis 487 Punkte)	Einfache grammatische Sprachbewusstheit: Eindeutige grammatische Fehler erkennen und z.T. korrigieren.	Im folgenden Satz, der aus einem Schüleraufsatz stammt, ist etwas grammatisch falsch. Unterstreiche die <u>grammatisch falsche Stelle</u> . <i>Die Diskothek wurde geschlossen, weil die Nachbarn die Lärmbelästigungen lange beklagt hatte.</i>

10.4 Ergebnisse

Verteilung auf die Kompetenzniveaus am Ende der neunten Jahrgangsstufe

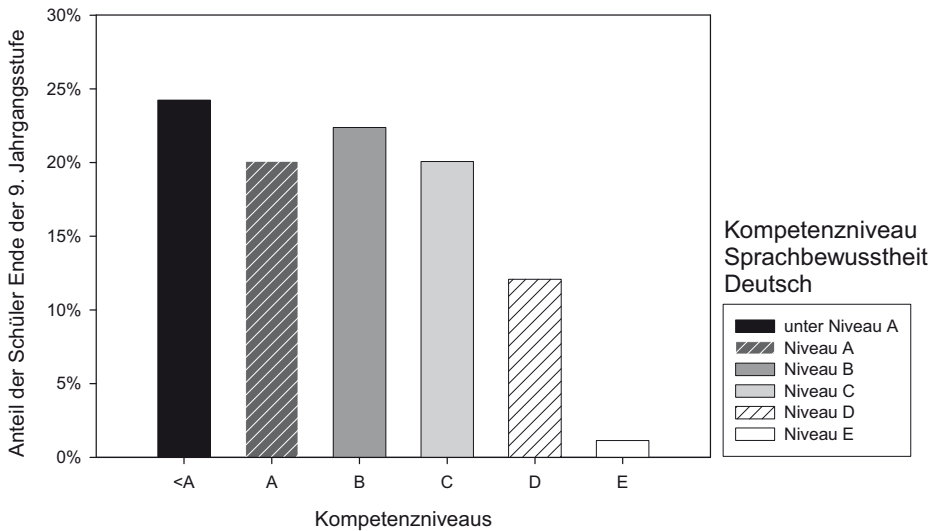


Abbildung 10.1: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Sprachbewusstheit Deutsch“ am Ende der neunten Jahrgangsstufe.

Die Kompetenzniveaus sind am Ende der neunten Jahrgangsstufe folgendermaßen besetzt (vgl. Abbildung 10.1):

- Kompetenzniveau A (20%): Diese Schülerinnen und Schüler können die grammatische Richtigkeit von Sprache erkennen und Korrekturen bei klaren grammatischen Verstößen (z.B. Dativ- Akkusativobjekt, Kongruenz, grammatisches Geschlecht) aus implizitem Wissen einbringen.
- Kompetenzniveau B (22%): Die Jugendlichen auf diesem Kompetenzniveau können zusätzlich Textkohärenz und -kohäsion aktiv erzeugen.
- Kompetenzniveau C (20%): Schülerinnen und Schüler auf diesem Niveau der Sprachbewusstheit können Korrekturen auch bei schwierigen, sprachwandelgefährdeten grammatischen Phänomenen wie z.B. dem Genitivobjekt anbringen. Sie korrigieren auch eindeutige Stilfehler. Darüber hinaus können sie grammatische Begriffe den entsprechenden Phänomenen zuordnen.
- Kompetenzniveau D (12%): Auf diesem Niveau ist die Fähigkeit, auch mit komplexeren sprachlichen und stilistischen Phänomenen korrekt umzugehen und grammatische Begriffe sicher zuzuordnen, weitgehend erworben.

- Kompetenzniveau E (1%): Auf diesem Niveau kommt noch die Fähigkeit hinzu, explizites (deklaratorisches) Wissen über komplexe Phänomene wie die indirekte Rede und die Formen des Konjunktivs aktiv zu nutzen, also z.B. die Funktion des Konjunktivs 2 zu benennen.

Die Schülerinnen und Schüler, die unter Kompetenzniveau A liegen (etwa noch 24%), sind im automatisierten Spracherwerb durchaus fortgeschritten. Sie sind jedoch in der sprachbewussten Ausarbeitung der zeitlich späteren Spracherwerbe im grammatischen System, z.B. sichere Unterscheidung von Dativ und Akkusativ, Genitivobjekt und hinsichtlich sprachwandelbedrohter Teilkompetenzen (z.B. Konjunktiv) noch sehr unsicher.

Verteilung nach Bildungsgängen

Interessant, wenn auch im Großen und Ganzen erwartungsgemäß, ist die Besetzung der Kompetenzniveaus nach Bildungsgängen (vgl. Abbildung 10.2). Bei der Verteilung am Ende der neunten Jahrgangsstufe zeigen sich ganz deutlich die großen Niveauunterschiede zwischen den Bildungsgängen.

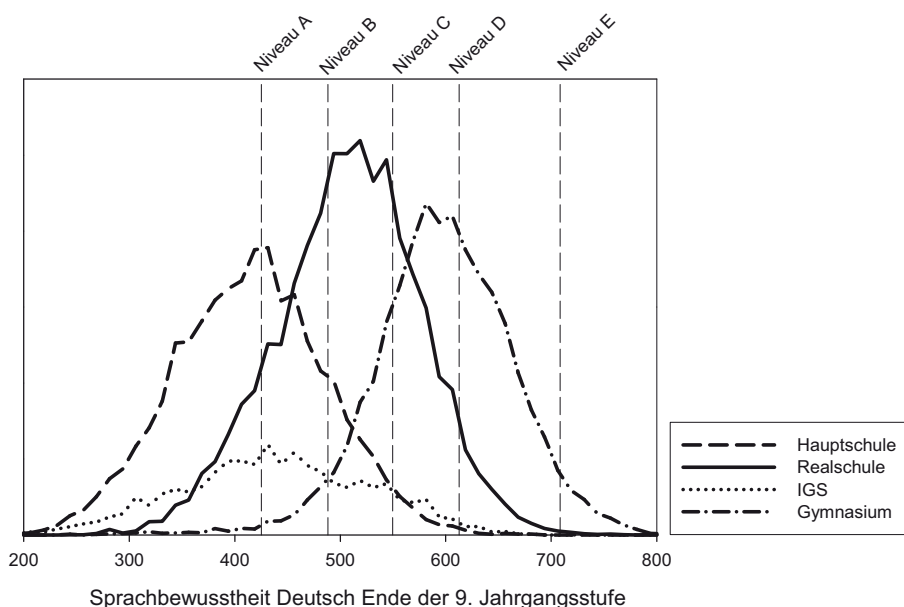


Abbildung 10.2: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Sprachbewusstheit Deutsch“ am Ende der neunten Jahrgangsstufe unter Beachtung der Bildungsgänge.

Sehr bedeutsam sind die großen Mittelwertsunterschiede zwischen den Bildungsgängen, die zum Teil zwei Kompetenzniveaus ausmachen sowie die nur geringen Überlappungen zwischen den Bildungsgängen.

Im Bereich Sprachbewusstheit zeigen sich am Ende der neunten Jahrgangsstufe so große Unterschiede zwischen Hauptschülern (Mittelwert 414 auf der DESI-Skala)

und Gymnasiasten (Mittelwert 598) wie bei keinem anderen Kompetenzbereich im Deutschen. Entscheidend ist, dass etwa 80% der Schülerinnen und Schüler im gymnasialen Bildungsgang, aber nur 3% der Hauptschüler über deklaratorisches Wissen in für dieses Schulalter relevanten Feldern verfügen, d.h. in der Lage sind, seltenere grammatische Phänomene zu benennen wie Konjunktiv und Indirekte Rede (Niveau C und höher). In den Integrierten Gesamtschulen liegt dieser Anteil bei 12%, im mittleren Bildungsgang bei 27%. Die Realschule steht „genau dazwischen“, während sich die IGS nur wenig besser als die Hauptschule präsentiert.

Zuwächse innerhalb der neunten Jahrgangsstufe

Abbildung 10.3 zeigt, dass in Sprachbewusstheit Deutsch ein außerordentlich deutlicher Zuwachs der Sprachleistungen zwischen den beiden Messzeitpunkten zu Beginn und am Ende der neunten Jahrgangsstufe zu verzeichnen ist.

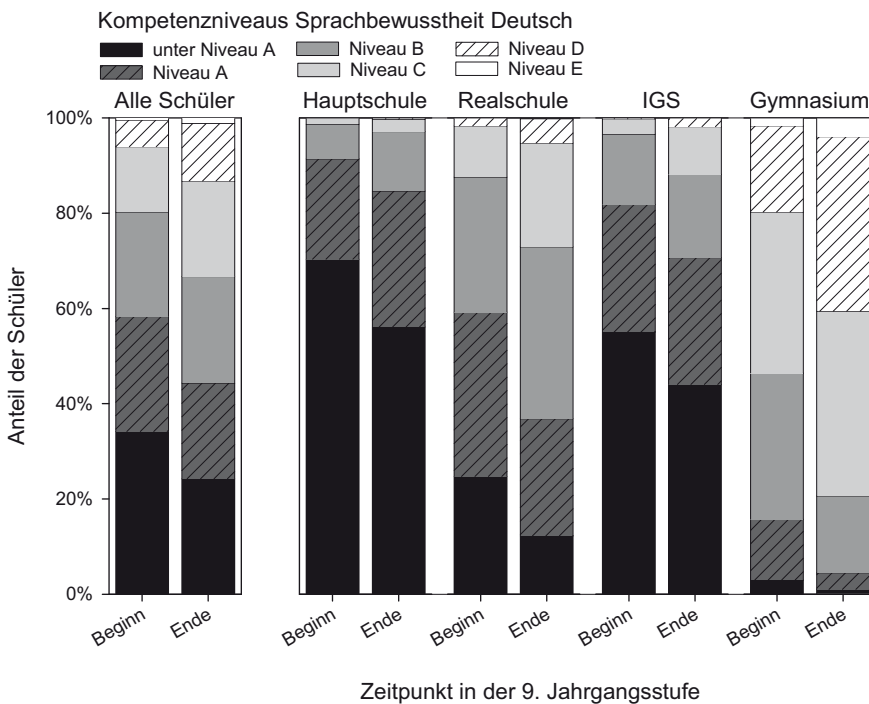


Abbildung 10.3: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in „Sprachbewusstheit Deutsch“ zu Beginn und am Ende der neunten Jahrgangsstufe unter Beachtung des Bildungsgangs.

Die Leistungszuwächse zwischen den beiden Messzeitpunkten zu Beginn und Ende der neunten Jahrgangsstufe sind in allen Bildungsgängen deutlich, sie reichen von 27 Punkten in Hauptschule und Integrierter Gesamtschule bis zu 36 Punkten in der Realschule und 42 Punkten im Gymnasium. Im Gymnasium – und nur hier wird in der neunten Jahrgangsstufe explizites grammatisches Wissen vermittelt – haben wir

Zuwächse vor allem in den beiden obersten Niveaus. Der Kompetenzzuwachs in der Realschule kommt dagegen vor allem dadurch zustande, dass dort der Anteil der Jugendlichen auf oder unter Niveau A von 59 auf 37 Prozent zurückgegangen ist. Der durchschnittliche Gewinn innerhalb der Jahrgangsstufe von knapp 35 Punkten liegt über den Zuwachsraten, die bisher bei Studien in der Sekundarstufe I in Deutschland gefunden wurden.

Die deutlichen Zuwächse als generelles Phänomen der gesamten Stichprobe kennzeichnen unseres Ermessens ein besonderes „Anwachsen“ der Sprachbewusstheit in der neunten Jahrgangsstufe, was sich am zunehmend besseren Umgang mit schwierigen grammatischen und stilistischen Phänomenen zeigt.

Leistungsverteilung nach Geschlechtern

Die Leistungsverteilung in Bewusstheit Deutsch nach Geschlechtern entspricht insofern unseren Erwartungen, als allgemein beobachtet und angenommen wird, dass die Mädchen den Jungen im Bereich sprachlicher Fähigkeiten überlegen sind (vgl. Kapitel 19). In allen Bildungsgängen und für die Stichprobe insgesamt wird das für Bewusstheit Deutsch deutlich dokumentiert.

Die Differenzen zwischen den Leistungen der Mädchen und Jungen sind zu beiden Testzeitpunkten in der IGS am deutlichsten (Beginn der neunten Jahrgangsstufe: Differenzwert 33, Ende der neunten Jahrgangsstufe: Differenzwert 55 zugunsten der Mädchen).¹ Das macht nicht selten ein Niveau aus. Im Gymnasium und in der Hauptschule liegen die Differenzwerte deutlich niedriger.

Die Zuwächse sind bei den Mädchen in allen Schularten ähnlich groß (Spanne zwischen 34 Leistungspunkten in der Hauptschule und 44 Leistungspunkten im Gymnasium). Die Realschule und die IGS liegen hier ganz dicht beim Gymnasium. Und diese Zuwächse sind in allen Schularten – bis auf das Gymnasium – deutlich größer als bei den Jungen, in der IGS sogar erheblich größer. Dieses Ergebnis legt nahe anzunehmen, dass die Mädchen vielleicht nicht nur eine höhere Sprachbewusstheit überhaupt haben, sondern dass sie sorgsamer sprachbewusst arbeiten als Jungen. Das ist ein wichtiges Signal, für eine auch kompensatorische Arbeit an der „sorgfältigen Sprache und am Stil“ im Unterricht der neunten und zehnten Jahrgangsstufe einzutreten.

Verteilung nach Sprachhintergrund

Die Schülerinnen und Schüler verhalten sich gemäß einer in der Literatur häufig anzutreffenden These, nämlich dass ein nicht deutscher Sprachhintergrund sich besonders in Defiziten bei den schwierigen grammatischen Phänomenen und in der Stilaneignung niederschlagen könnte (vgl. Kapitel 20). Die Schülerinnen und Schüler nicht deutscher Erstsprache befinden sich noch deutlich im Stadium des Erwerbs der sprachbewussten System- und Stilkompetenzen, was die Differenzen in

¹ Die Zahlenangaben beziehen sich auf Punkte der DESI-Skala.

den Leistungspunkten vor allem der Hauptschule und der IGS deutlich belegen. Die Gruppe der „Mehrsprachigen“ liegt in ihren Leistungen ziemlich genau zwischen der Gruppe mit Deutsch als Erstsprache und der Gruppe der Schülerinnen und Schüler nicht deutscher Erstsprache.

Bei den Zuwächsen zeigt sich ein besonders interessantes Ergebnis: Die Schülerinnen und Schüler nicht deutscher Erstsprache weisen deutlich größere Leistungszuwächse auf als die deutschsprachigen Schülerinnen und Schüler. Eine Ausnahme bilden die Ergebnisse der IGS, dort ist die Relation genau umgekehrt. Noch eine andere interessante Beobachtung gilt es mitzuteilen: Die mehrsprachigen Schülerinnen und Schüler des Bildungsganges Gymnasium sind im Durchschnitt besser als die Deutschsprachigen in den anderen Bildungsgängen.

Literatur

- Andresen, H./Funk, R. (2003): Entwicklung sprachlichen Wissens und sprachlicher Bewusstheit. In: Bredel, U./ Günther, H./Klotz, P./Ossner, J./Siebert-Ott, G. (Hrsg.): Didaktik der Deutschen Sprache. Band 1. Paderborn, S. 438-451.
- Eichler, W. (2003): Die Pisa-Nachfolgestudie DESI, Deutsch-Englische Sprachkompetenz bei SchülerInnen des 9. Schuljahrs international. In: Moschner, B./Kiper, H./Kattmann, U. (Hrsg.): PISA 2000 als Herausforderung. Perspektiven für Lehren und Lernen. Baltmannsweiler: Schneider, S. 157-172.
- Eichler, W. (2004a): Sprachbewusstheit und grammatisches Wissen. In: Grundschule 10, S. 58-61.
- Eichler, W. (2004b): Sprachbewusstheit und Orthographieerwerb. In: Bremerich-Vos, A./Löffler, C./Herné, K.-L. (Hrsg.): Neue Beiträge zur Rechtschreibtheorie und -didaktik, Freiburg i. Brsg.: Fillibach, S. 179-189.
- Eichler, W. (2007a): Kompetenzmodelle und Kompetenzniveaus im Bereich des Deutschen: Sprachbewusstheit. In: Beck B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 147-157.
- Eichler, W. (2007b): Sprachbewusstheit in DESI. In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Baltmannsweiler: Schneider, S. 124-133.
- Eichler, W./Nold, E. (2007): Übergreifende Konzeptualisierung sprachlicher Kompetenzen. Sprachbewusstheit. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band I. Weinheim: Beltz, S. 63-82.
- Karmiloff-Smith, A. (1992): Beyond Modularity: A Developmental Perspective on Cognitive Science. **Boston: MIT Press.**
- Luchtenberg, S. (1995): Language Awareness-Konzeptionen. In: Der Deutschunterricht 4, S. 3-14.

Günter Nold / Henning Rossa

11 Hörverstehen Englisch

11.1 Testkonzeption

Im Einklang mit dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GERS) des Europarats wird Hörverstehen in DESI als eine rezeptive kommunikative Aktivität erfasst, die mit Verstehensabsichten verknüpft ist. So können Zuhörerinnen und Zuhörer beispielsweise beabsichtigen zu erfahren, was in einem Gespräch insgesamt gemeint ist, oder sie möchten eine ganz bestimmte Information erhalten oder das Gesprochene in allen Einzelheiten verstehen, um Schlussfolgerungen ziehen zu können (vgl. Europarat 2001). Die Verstehensabsichten finden sich in der Form von Lernzielen, Strategien und Aufgabenstellungen auch in den curricularen Vorgaben zur Entwicklung des Hörverstehens im Fremdsprachenunterricht (vgl. Nold/Rossa 2007) und bilden zusammen mit den entsprechenden Textsorten, Diskurstypen und Themen den Rahmen, in dem die Lernenden ihre Kompetenzen einsetzen müssen, um die kommunikative Aktivität des Hörverstehens erfolgreich zu gestalten. Beim Hörverstehen in der Fremdsprache Englisch benötigen die Lernenden neben Weltwissen und strategischen Kompetenzen, die die Verarbeitung von sprachlichen Informationen erlauben, vor allem schnell verfügbares Sprachwissen, um trotz der Flüchtigkeit der Informationen ein Verständnis des Gehörten zu entwickeln (vgl. Nold/Rossa 2007). Das Testkonstrukt des fremdsprachlichen Hörverstehens Englisch spezifiziert die Bedingungen, unter denen Aspekte der Kompetenz im Bereich Hörverstehen durch die Aufgaben erfasst werden sollen.

In DESI bezieht sich die Kompetenz des Hörverstehens auf die Fähigkeit, kurze und etwas umfangreichere Beispiele gesprochener Sprache zu verarbeiten, um Details und Hauptaussagen sowie implizite Äußerungen zu verstehen. Die Schülerinnen und Schüler haben zu Beginn und am Ende der neunten Klasse jeweils zwei Gruppen von Hörverstehensaufgaben bearbeitet. Zur ersten Gruppe gehören 16 kurze Dialoge (10 bis 20 Sekunden), zu denen jeweils eine Aufgabe vom Typ *Multiple-Choice* gestellt wurde. Die zweite Gruppe besteht aus längeren Texten (ca. zwei Minuten), die einem Radiobeitrag ähneln. Zu diesen Texten waren jeweils sieben bis zehn *Multiple-Choice*-Aufgaben zu bearbeiten. Die Hörtexte präsentieren eine jugendliche Sprecherin und einen erwachsenen Sprecher, deren Varietäten des Englischen *General Canadian*¹ beziehungsweise *Near-Received Pronunciation* sind. Bei meist deutlicher Artikulation ist die Sprechgeschwindigkeit der Dialoge,

1 Es wurden Akzente des Englischen gewählt, die den Zugang der Schülerinnen und Schüler zu den gesprochenen Texten nicht übermäßig erschweren sollten, da sie vorwiegend an Sprachvarianten des britischen Englisch gewöhnt sind. Der Akzent *General Canadian* stellt in diesem Fall im Vergleich zu US-amerikanischen Akzenten einen milden Kontrast zu britischen Standardakzenten dar.

Berichte und Erzählungen annähernd normal. Es wird allerdings nicht immer gleich schnell gesprochen.

11.2 Kompetenzniveaus und Aufgabenbeispiele

Die Hörtexte und die dazugehörigen Aufgaben reflektieren in quantitativ und qualitativ unterschiedlichen Ausprägungen Elemente des Testkonstrukts. Diese Ausprägungen bestimmen die Anforderungen, die von den Schülerinnen und Schülern beim Bearbeiten der Aufgaben zu bewältigen sind. In DESI wurden die Anforderungsprofile der Aufgaben im Zusammenhang mit der Aufgabenschwierigkeit analysiert und zur Markierung von Kompetenzniveaus auf einer kontinuierlichen Skala genutzt. Je nachdem, welche Aufgaben eine Person sicher zu lösen im Stande ist, wird sie in DESI einem der drei Kompetenzniveaus der Skala Hörverstehen Englisch zugeordnet. Die charakteristischen Merkmale der Aufgaben, die die Anforderungen der Niveaus bestimmen, wurden bei der Entwicklung der Skala in Kann-Beschreibungen ähnlich denen der Referenzniveaus des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GERS) umgewandelt (vgl. Tabelle 11.1) (zur Entwicklung des Kompetenzmodells Hörverstehen in DESI vgl. Nold/Rossa 2007).

Auf dem untersten Kompetenzniveau A können die Schülerinnen und Schüler konkrete Einzelinformationen aus Kontexten alltäglicher Kommunikation verstehen, wenn diese Informationen mehr oder weniger wörtlich im Text enthalten sind, relativ langsam sowie deutlich gesprochen und in einfacher Sprache präsentiert werden. Neben diesem elementaren Detailverstehen sind sie in der Lage, eine sehr begrenzte Anzahl von Informationen beim Hören zu verknüpfen, um Hauptaussagen zu verstehen.

Schülerinnen und Schüler, deren Testleistung dem mittleren Kompetenzniveau B zugeordnet wurde, können zusätzlich zu der auf Kompetenzniveau A beschriebenen elementaren Kompetenz implizit geäußerte Informationen erschließen (schlussfolgern) und wörtlich aus dem Text entnommene Informationen interpretieren. Auf diesem Niveau werden zudem abstraktere Informationen in alltäglichen Kontexten (z.B. Äußerungen über Emotionen) verstanden, auch wenn die Äußerungen sprachlich komplexer sind und in normaler Sprechgeschwindigkeit präsentiert werden.

Das höchste Kompetenzniveau C erreichen Schülerinnen und Schüler, die abstrakte Informationen ohne direkten Alltagsbezug (z.B. Gegensätze, Unterscheidungen, Textstrukturen) verstehen, auch wenn diese sprachlich komplex und in partiell schneller Sprechgeschwindigkeit präsentiert werden, wie Muttersprachler dies in natürlicher Interaktion tun. Außerdem können Personen auf diesem Niveau mehrere Informationen beim Hören verknüpfen, um die Hauptaussage zu verstehen, auch wenn diese Informationen über eine längere Äußerung hinweg verteilt sind.

Tabelle 11.1: Beschreibung der Kompetenzniveaus in Hörverstehen Englisch mit Aufgabenbeispielen.

Kompetenz-niveau	Beschreibung
C+ C (ab 651 Punkten)	<p>Kann mehrere Informationen beim Hören verknüpfen, um die Hauptaussage zu verstehen, auch wenn diese Informationen über eine längere Äußerung hinweg verteilt sind.</p> <p>Kann abstrakte Informationen ohne direkten Alltagsbezug (z.B. Gegensätze, Unterscheidungen, Textstrukturen) verstehen, indem implizite Informationen erschlossen oder inhaltlich komplexe Einzelinformationen interpretiert werden, auch wenn diese sprachlich komplex und in partiell schneller Sprechgeschwindigkeit präsentiert werden, wie Muttersprachler dies in natürlicher Interaktion tun.</p>
B+ B (541 bis 650 Punkte)	<p>Kann abstraktere Informationen in alltäglichen Kontexten (z.B. Äußerungen über Emotionen) verstehen, auch wenn diese sprachlich komplex und in normaler Sprechgeschwindigkeit präsentiert werden.</p> <p>Kann ein Verständnis von konkreten Informationen entwickeln, indem implizite Informationen erschlossen (Schlussfolgern) oder explizite Informationen interpretiert werden.</p>
A+ A (451 bis 540 Punkte)	<p>Kann konkrete Einzelinformationen aus Kontexten alltäglicher Kommunikation (Erzählungen, Radioberichte, Gespräche) hörend verstehen und eindeutige Umschreibungen dieser Informationen erschließen, auch wenn diese mit einer etwas breiteren Auswahl sprachlicher Mittel sowie in normaler Sprechgeschwindigkeit präsentiert werden. Kann eine geringe Anzahl konkreter Informationen beim Hören verknüpfen, um Hauptaussagen zu verstehen.</p> <p>Kann konkrete Einzelinformationen aus Kontexten alltäglicher Kommunikation (Erzählungen, Radioberichte, Gespräche) hörend verstehen, wenn diese Informationen langsam und deutlich gesprochen und in einfacher Sprache explizit präsentiert werden.</p>

Aufgabenbeispiele

In den folgenden Abbildungen 11.1 bis 11.4 werden Aufgabenbeispiele für die drei Kompetenzniveaus dargestellt. Eine detaillierte Darstellung findet sich in Nold/Rossa (2007). Bei Testaufgaben stellt sich unter anderem die Frage, welche Rolle das Weltwissen bei der Beantwortung der Verstehensfrage spielt. Da erfolgreiches Hörverstehen jedoch stärker von der Fähigkeit automatisierter Informationsverarbeitung abhängig ist als von der Aktivierung bereichsspezifischen Weltwissens (vgl. Kintsch 1998; Buck 2001), kommt der Frage, ob die Testteilnehmer mit dem Inhalt beispielsweise von Tolkiens *Lord of the Rings* vertraut sind, auf einem höheren Niveau eine untergeordnete Rolle zu (vgl. die ausführliche Darstellung in Nold/Rossa 2007).

[...]
 Woman: You were supposed to meet me at the entrance at *three*.
 Man: What? I thought we had agreed to meet at three-thirty.
 Woman: Well, it's four now so that's no excuse. I just can't count on you. You're always late.
 [...]

		Why is the woman angry with the man?
A	<input type="checkbox"/>	He was not at the entrance at four.
B	<input checked="" type="checkbox"/>	He has been late many times.
C	<input type="checkbox"/>	He didn't wait for her.

Abbildung 11.1: Aufgabenbeispiel für Kompetenzniveau A in Hörverstehen Englisch.

Woman: Wow! Where did you get that CD from? Do shops already have it?
 Man: My brother works at a record store in the city and he gave it to me on my birthday last week.
 [...]

		How did the man get the CD?
A	<input type="checkbox"/>	He bought it at a store.
B	<input checked="" type="checkbox"/>	He got it as a present.
C	<input type="checkbox"/>	He took it from his brother.

Abbildung 11.2: Aufgabenbeispiel für Kompetenzniveau A+ in Hörverstehen Englisch.

Man: I'm not really looking forward to that party.
 Woman: You aren't?
 Man: Well, I've been to parties like this one before, and I just think that – there will be too many people who go on and on about how wonderful they are, and they think they're so important. [...] they, erm, well, they make me feel uneasy. I never know what to say.
 [...]

		How does the man feel about the party?
A	<input checked="" type="checkbox"/>	He is nervous.
B	<input type="checkbox"/>	He doesn't know what to expect.
C	<input type="checkbox"/>	He thinks it will be boring.

Abbildung 11.3: Aufgabenbeispiel für Kompetenzniveau B in Hörverstehen Englisch.

A language expert

David Salo is an expert in languages that exist only in the imaginary world of hobbits and elves. He has immersed himself in Quenya and Sindarin, languages created by author J.R.R. Tolkien for the inhabitants of Middle-earth and featured in his Lord of the Rings trilogy.

So when filmmakers adapting the fantasy epic wanted to translate parts of their script from English into the two Elvish languages, they turned to Salo, a graduate student in linguistics at the University of Wisconsin-Madison.

[...]

"I found the script in my mailbox two weeks later, and it was only then that I realised I was really involved in the making of the movie."

For Salo, a die-hard fan of Tolkien's books, this must have been a dream come true.

		What is the topic of this radio report?
A	<input type="checkbox"/>	David Salo's life.
B	<input checked="" type="checkbox"/>	How a student was able to use his special language skills.
C	<input type="checkbox"/>	The movie version of The Lord of the Rings.

Abbildung 11.4: Aufgabenbeispiel für Kompetenzniveau C in Hörverstehen Englisch.

11.3 Verteilung auf die Kompetenzniveaus

Im Folgenden wird berichtet, wie sich die Testleistungen der Schülerinnen und Schüler am Ende der neunten Klasse auf die Kompetenzniveaus verteilen. Wie Abbildung 11.5 zeigt, verteilt sich knapp ein Drittel der Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe auf das Kompetenzniveau A und verfügt somit mindestens über eine elementare Kompetenz im Bereich Hörverstehen. 28% der Schülerinnen und Schüler sind in der Lage, auch bei normaler Sprechgeschwindigkeit implizit Geäußertes zu verstehen und Informationen zu interpretieren (Kompetenzniveau B).

Ein gutes Drittel der Neuntklässler löst jedoch die Aufgaben des Kompetenzniveaus A nicht mit ausreichender Sicherheit. Es ist dennoch davon auszugehen, dass viele dieser Schülerinnen und Schüler mit einer gewissen Wahrscheinlichkeit Details in kürzeren Äußerungen verstehen können. Insgesamt ist das Verständnis aber oft so lückenhaft, dass es den Schülerinnen und Schülern nicht gelingt, Details und auch einige Hauptaussagen so zu erfassen, wie es die Aufgaben auf Kompetenzniveau A erfordern.

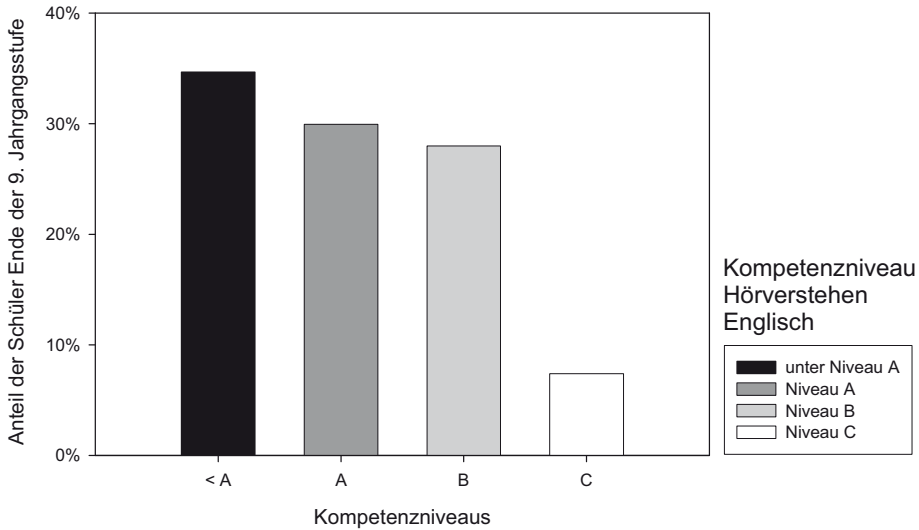


Abbildung 11.5: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Hörverstehen Englisch am Ende der neunten Jahrgangsstufe.

Die Schülerinnen und Schüler, deren Testleistungen unter das Niveau A eingestuft wurden, können dementsprechend die curricularen Mindestanforderungen, die sich an den Deskriptoren des Niveaus A2 des Gemeinsamen Europäischen Referenzrahmens orientieren, nur bruchstückhaft bewältigen. Andererseits haben 7% der Neuntklässler eine Kompetenz entwickelt, die es ihnen erlaubt, auch abstrakte und sprachlich komplexe Informationen zu verstehen, die in natürlicher Sprechgeschwindigkeit vorgelesen werden (Kompetenzniveau C). Sie erreichen damit ein Niveau, das deutlich über den curricularen Mindestanforderungen liegt.

11.4 Verteilung auf die Kompetenzniveaus und Lernzuwachs: Unterschiede zwischen den Bildungsgängen

Im Vergleich der Bildungsgänge fällt auf, dass sich die Schülerinnen und Schüler in jedem Bildungsgang – wenn auch in stark unterschiedlichen Anteilen – auf alle Kompetenzniveaus verteilen (vgl. Abbildung 11.6). Auf dem höchsten Niveau variiert die Verteilung zwischen 21% der Schülerinnen und Schüler am Gymnasium und 0.2% an Hauptschulen. Zugleich gibt es in allen Bildungsgängen Schülerinnen und Schüler, deren Kompetenz nicht ausreicht, um die Aufgaben des ersten Kompetenzniveaus sicher lösen zu können. Die Extremwerte bewegen sich hier zwischen 3% und 70% der Schülerinnen und Schüler eines Bildungsganges.

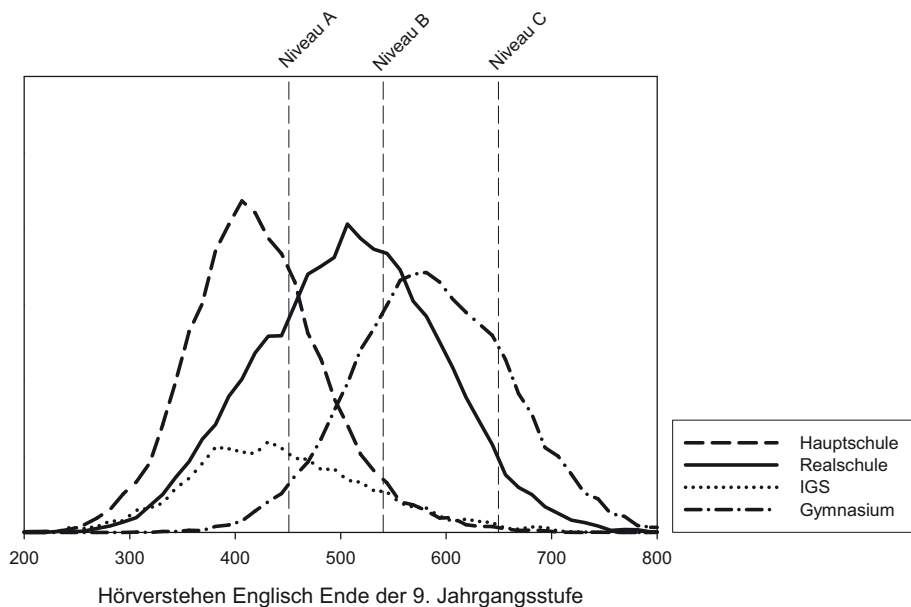


Abbildung 11.6: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Hörverstehen Englisch am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Da die Schülerinnen und Schüler sowohl am Anfang als auch am Ende der neunten Jahrgangsstufe Aufgaben zum Hörverstehen bearbeitet haben, ist es möglich die Veränderungen in der Kompetenzentwicklung abzuschätzen. Wie Abbildung 11.7 erkennen lässt, ist am Ende der neunten Klasse ein deutlicher Lernzuwachs festzustellen. So steigt der Anteil der Schülerinnen und Schüler auf den Niveaus B und C von 25% auf 35%. Gleichzeitig sinkt der Anteil der Schülerinnen und Schüler auf den Niveaus A und unter A entsprechend von 75% auf 65%.

Am Ende der neunten Klasse liegt die Kompetenz im Bereich des Hörverstehens bei Schülerinnen und Schülern der Hauptschule im Durchschnitt ($M = 422$) deutlich, an der Gesamtschule ($M = 445$) knapp unter Kompetenzniveau A. Schülerinnen und Schüler der Realschule erreichen durchschnittlich das Kompetenzniveau A ($M = 508$) und am Gymnasium das Kompetenzniveau B ($M = 587$). Die durchschnittlichen Testleistungen verbessern sich im Verlauf einer Jahrgangsstufe in den Bildungsgängen Hauptschule und Realschule um 5% und auf der Gesamtschule und am Gymnasium um 6%. Abbildung 11.7 zeigt, wie sich dieser Lernzuwachs je nach Bildungsgang in recht unterschiedlichen Zugangs- und Abgangsraten der einzelnen Kompetenzniveaus offenbart.

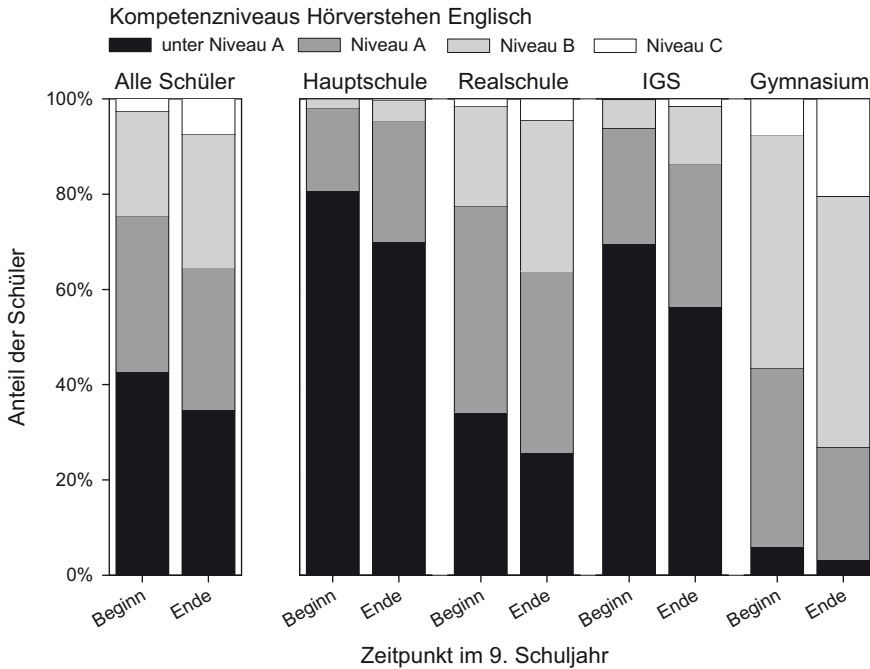


Abbildung 11.7: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Hörverstehen Englisch zu Beginn und am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Im Bildungsgang der Hauptschule sinkt der Anteil der Schülerinnen und Schüler, die unter das Kompetenzniveau A eingestuft werden, um 11%. Den größten Zuwachs erfährt das Kompetenzniveau A mit 8%. In der Realschule ist der größte Zuwachs (11%) auf Kompetenzniveau B zu verzeichnen. Der Anteil der Schülerinnen und Schüler, die im Bildungsgang der Gesamtschule unter das Kompetenzniveau A eingestuft wurden, sinkt um 13%, während jeweils 6% mehr Schülerinnen und Schüler auf die Kompetenzniveaus A und B eingestuft werden. Am Gymnasium verzeichnet das höchste Kompetenzniveau den größten Zuwachs: Der Anteil der Schülerinnen und Schüler auf diesem Niveau steigt von 8% auf 21%.

11.5 Diskussion

Ein Vierteljahrhundert nach der kommunikativen Wende im Fremdsprachenunterricht scheint die Hinwendung zur Mündlichkeit mit dem Slogan „fluency before accuracy“ (vgl. Lightbown/Spada 2006) in deutschen Schulen zumindest für die Entwicklung von Kompetenz im Hörverstehen bedingt Früchte zu tragen. Nach fünf Jahren Englischunterricht reichen die Fähigkeiten eines Drittels der Schülerinnen und Schüler nicht dazu aus, konkrete Einzelinformationen in Dialogen und Berichten zu verstehen, die in relativ langsamer Sprechgeschwindigkeit und deutlicher Artikulation präsentiert werden. Um diese Gruppe von Lernenden zu fördern, müsste der Englischunterricht

in der Sekundarstufe I mehr Lerngelegenheiten schaffen, die den schnellen Abruf des sich entwickelnden Sprachwissens und -könnens trainieren, beispielsweise durch den regelmäßigen Einsatz kurzer Hörtexte in der Wortschatzvermittlung oder zum Einstieg in ein neues Thema. Zusätzlich kann ein Training von Wortschließungsstrategien den Lernenden helfen, im Hörverstehensprozess fehlendes oder nur in geringem Maße automatisiertes Sprachwissen zu kompensieren (vgl. Mendelsohn 1995).

Es ist positiv zu bewerten, dass eine große Gruppe von Schülerinnen und Schülern im Englischunterricht offenbar Kompetenzen entwickelt, die über die curricular angestrebten Ziele hinausgehen. Sie profitieren von der kommunikativen Ausrichtung des Unterrichts und verfügen am Ende der neunten Klasse über automatisiertes Sprachwissen, das sie für Hörverstehen auf einem fortgeschrittenen Niveau einsetzen können. Die Fähigkeit, ein relativ stabiles Verständnis des Gehörten zu konstruieren, versetzt sie zudem in die Lage, den Verstehensprozess durch Schlussfolgern zu vertiefen und für weitergehende Interpretationen zu öffnen.

Für die Zukunft lassen gegenwärtige Entwicklungen die Hoffnung zu, dass das Hörverstehen auch in der Sekundarstufe I als eine kommunikative Aktivität begriffen wird, die für den Erwerb der Fremdsprache unentbehrlich ist und in der Unterrichtsgestaltung entsprechend regelmäßig zu berücksichtigen ist. Das Prinzip der Mündlichkeit im Fremdsprachenunterricht in der Primarstufe kann beispielsweise dazu führen, dass als Ergebnis der Diskussion um die Gestaltung des Übergangs zum Englischunterricht in der Sekundarstufe die Rolle des Hörverstehens im Unterricht der Klassen fünf bis zehn stärker betont wird. Das mediale Potential an authentischen Texten auf CD-Rom oder DVD lässt sich zudem im Unterricht zeitökonomisch und leicht handhabbar nutzen. Auch die Vorboten der neuen Evaluationskultur – zentrale Prüfungen, Lernstandserhebungen, einheitliche Prüfungsanforderungen etc. (vgl. Rossa 2006) besitzen bei Berücksichtigung des Hörverstehens als Prüfungselement das Potenzial, die Beachtung der mündlichen Kompetenzen bei Lehrenden sowie Schülerinnen und Schülern zu stärken.

Literatur

- Beck, B./Klieme, E. (Hrsg.) (2007): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Buck, G. (2001): Assessing Listening. Cambridge: CUP.
- Europarat (Hrsg.) (2001): Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen. Berlin: Langenscheidt.
- Kintsch, W. (1998): Comprehension. Cambridge: Cambridge University Press.
- Lightbown, P./Spada, N. (2006): How languages are learned. Third Edition. Oxford: Oxford University Press.
- Mendelsohn, D.J. (1995): Applying Learning Strategies in the Second/Foreign Language Listening Comprehension Lesson. In: Mendelsohn, D. J./Rubin, J. (Hrsg.): A Guide for the Teaching of Second Language Listening. Carlsbad, CA: Dominie Press, S. 132-150.
- Mendelsohn, D. J./Rubin, J. (Hrsg.) (1995): A Guide for the Teaching of Second Language Listening. Carlsbad, CA: Dominie Press.

- Nold, G./Rossa, H. (2007): Hörverstehen. In: Beck, B. & Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 174-192.
- Rossa, H. (2006): Large-scale assesement in DESI: Wie lassen sich Kompetenzen in der Fremdsprache Englisch erfassen? In: Praxis Fremdsprachenunterricht 2, S. 2-6.

Günter Nold / Henning Rossa / Kyriaki Chatzivassiliadou

12 Leseverstehen Englisch

12.1 Testkonzeption

Leseverstehen beruht auf dem Zusammenspiel verschiedener Wissensbestände und Fähigkeiten (vgl. Alderson 2000; Artelt u.a. 2001; Grotjahn 2000; Kintsch 1998; Urquhart/Weir 1998; Smith 1971). Einerseits werden sprachliche Zeichen entschlüsselt und Bedeutungen erschlossen, andererseits werden Informationen mit Weltwissen verknüpft. Im Leseprozess wird dabei die Aufmerksamkeit entsprechend den mit dem Lesen verbundenen Absichten fokussiert und strategisch in unterschiedlichen Lesearten (beispielsweise *skimming*, *scanning*, *selective reading*) umgesetzt. Aus diesen grundsätzlichen Überlegungen lassen sich für die Testkonstruktion spezifische Aspekte entwickeln, die festlegen, was konkret zum Gegenstand eines Tests gemacht wird. In den DESI-Tests zum Leseverstehen spielen zusätzlich zu diesen theoretischen Überlegungen curriculare Konzeptionen der Bundesländer und Vorstellungen des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GERS; Europarat 2001; Council of Europe 2001) eine Rolle bei der Festlegung des Testkonstrukts. So wird Leseverstehen in DESI als eine Fähigkeit zur Teilhabe an kommunikativen Aktivitäten verstanden (vgl. Europarat 2001), wobei angenommen wird, dass sich diese Fähigkeit im Kontext des Fremdsprachenunterrichts und in außerschulischen Interessensfeldern entwickelt.

Im Einzelnen geht es im Testkonstrukt Leseverstehen um die Fähigkeit, explizit und implizit präsentierte Informationen (z.B. Ereignisse, Emotionen, Meinungen) mit Hilfe von Sprachwissen zu erkennen und zu erschließen. Um diese Informationsentnahme und Bedeutungszuweisung leisten zu können, ist es vielfach erforderlich, dass der Leser im Text angelegte Informationen mit curricular voraussetzbarem Weltwissen verknüpft und somit Informationen selbst erkennt, erschließt und in einer Zusammenschau interpretiert. Der Fremdsprachenlerner benötigt außerdem die Fähigkeit, unbekannte sprachliche Elemente aus dem Kontext erschließen zu können, um Hauptaussagen zu verstehen und die Kohärenz des Gesamttextes und von Textteilen herzustellen.

Die genannten Facetten des Leseverstehens werden in Aufgaben erfasst, die sich auf narrative Texte, Berichte, Erzählungen und Dramentexte mit unterschiedlichen sprachlichen und textpragmatischen Niveaus beziehen (vgl. Nold/Rossa 2006, 2007; Nold/Willenberg 2007).

12.2 Kompetenzniveaus und Aufgabenbeispiele

Die Kompetenzniveaus zum Leseverstehen Englisch bestehen aus gestuften Kann-Beschreibungen, die auf der Grundlage von Aufgabenmerkmalsbeschreibungen

der Test-Items und den empirischen Schwierigkeitswerten der Lesetestergebnisse entwickelt wurden (vgl. Nold/Rossa 2007; Nold/Rossa/Hartig in Druck). Als unterstes Niveau wurde dabei der curricular angestrebte Mindeststandard zu Grunde gelegt. Die verschiedenen Niveaus der Lesekompetenz implizieren, dass ein Testteilnehmer, dessen Testleistung einem Niveau zugeordnet wird, mit einer hohen Wahrscheinlichkeit eine Kompetenz entwickelt hat, die es ihm erlaubt, unter den beschriebenen Bedingungen erfolgreich in der Fremdsprache Englisch lesen zu können. Jede Niveaubeschreibung schließt dabei die Beschreibungen der darunter liegenden Niveaus mit ein (vgl. Tabelle 12.1). Ursprünglich (vgl. Nold/Rossa 2007) wurde zwischen vier Kompetenzniveaus des Leseverstehens (A, B, C, D) unterschieden. Aufgrund der engen Lage der empirisch ermittelten Schwellen zwischen den Niveaus wurden die ersten beiden Niveaus A und B zu einem Niveau A zusammengefasst. Die ursprüngliche inhaltliche Differenzierung der Anforderungen innerhalb dieses Niveaus werden durch ein Niveau A+ innerhalb von Niveau A beschrieben.

Tabelle 12.1: Beschreibung der Kompetenzniveaus in Leseverstehen Englisch.

Kompetenz-niveau	Beschreibung
C (ab 619 Punkte)	Kann abstrakte Informationen (z.B. Meinungen, Textstrukturen) mit Hilfe von Inferieren impliziter Informationen verknüpfen oder inhaltlich komplexe Einzelinformationen interpretieren, auch wenn die Texte insgesamt einen erweiterten Wortschatz und eine begrenzte Anzahl von komplexen Strukturen enthalten sowie wenige textverknüpfende Elemente enthalten.
B (582 bis 618 Punkte)	Kann eine begrenzte Anzahl abstrakterer Informationen (z.B. Emotionen) verknüpfen und interpretieren, um Hauptaussagen zu verstehen, wenn die Texte Grundwortschatz, weitgehend gebräuchliche Strukturen sowie deutlich textverknüpfende Elemente enthalten.
A+ (498 bis 581 Punkte)	Kann abstraktere Einzelinformationen (z.B. Emotionen) in alltäglichen Kontexten erkennen, auch wenn einzelne Textpassagen weniger frequente Wörter und einige komplexere Strukturen enthalten.
A (475 bis 581 Punkte)	Kann konkrete Einzelinformationen in alltäglichen Kontexten (narrative Texte/Berichte) anhand von expliziten Hinweisen im Text (Schlüsselwörter und Umschreibungen) erkennen, wenn die fokussierten Textteile im Wesentlichen in einfacher Sprache und inhaltlich deutlich kohärent abgefasst sind.

Aufgabenbeispiele

In den folgenden Abbildungen 12.1 bis 12.3 werden Aufgabenbeispiele für die drei Kompetenzniveaus dargestellt. Eine detaillierte Darstellung findet sich in Nold/Rossa (2007).

Arrested

In Washington, DC, a policeman arrested and handcuffed a 14-year-old girl because she was eating a cheeseburger in a Metro underground station. Does that sound too strange to be true? It really happened to Anne Hedgepeth.

[...]

Ray Feldman, a Metro representative, said: "We had to take action because it's illegal to eat and drink in the Metro system, and we are tired of people who eat and drink on the train, spill things, and leave food and cans."

[...]

Hedgepeth's mother later wrote a letter to complain to Mr Hewitt. He answered: "I am sorry that it was necessary to arrest your daughter. I hope you understand the responsibility we have to keep public transport safe and clean."

Kompetenzniveau A

		Anne Hedgepeth was eating a cheeseburger
A	<input type="checkbox"/>	at a bus station.
B	<input type="checkbox"/>	in front of the police station.
C	<input checked="" type="checkbox"/>	in an underground station.
D	<input type="checkbox"/>	on the train.

Kompetenzniveau A+

		The police arrest people who eat and drink in the Metro because
A	<input type="checkbox"/>	they think it is dangerous.
B	<input type="checkbox"/>	mothers complain about the mess.
C	<input checked="" type="checkbox"/>	they want to keep the trains clean.
D	<input type="checkbox"/>	they think it is unhealthy.

Abbildung 12.1: Aufgabenbeispiele für Kompetenzniveaus A und A+ in Leseverstehen Englisch.

Alicia Keys

She's 20 years old and she's the new voice of R&B soul. Her album *Songs in a Minor* was number one in the US charts. Her single *Fallin'* was number one in the single charts. But Alicia Keys isn't falling. She's rising, and she's rising fast.

How did beautiful Alicia get so big so fast? First, she started writing some of the songs from her album when she was 14. When she was sixteen she already had her high school diploma and was going to Columbia University in her home city, New York. "I was so deep into music, I didn't want to be in high school cliques and gossip," Alicia explains. But soon after she got to university she left because she knew she wanted to concentrate on music. [...]

		Alicia became big very fast because
A	<input type="checkbox"/>	she went to Columbia University.
B	<input type="checkbox"/>	she lived in New York.
C	<input checked="" type="checkbox"/>	she concentrated completely on music.
D	<input type="checkbox"/>	she was in high school cliques.

Abbildung 12.2: Aufgabenbeispiel für Kompetenzniveau B in Leseverstehen Englisch.

A comfortless flat in Manchester.

Enter Helen and Jane, loaded with baggage.

HELEN: Well, this is the place.
JANE: And I don't like it.
HELEN: When I find somewhere for us to live I have to consider far more important things than your feelings... the rent. It's all I can afford.

[...]
JANE: Can't be soon enough for me. I'm cold and my shoes let water ... what a place ... and I'm supposed to be living off her earnings.

HELEN: I'm careful. Anyway, what's wrong with this place? Everything in it is falling apart, it's true, but there's a lovely view of the gasworks, we share a bathroom with the community and this wallpaper's contemporary. What more do you want?

(from Top Girls, by Caryl Churchill)

		It seems that Helen and Jane
A	<input type="checkbox"/>	don't mind living in this flat.
B	<input checked="" type="checkbox"/>	are not very happy to be in this place.
C	<input type="checkbox"/>	have lived in better places before.
D	<input type="checkbox"/>	like the flat because of its nice view.

Abbildung 12.3: Aufgabenbeispiel für Kompetenzniveau C in Leseverstehen Englisch.

12.3 Verteilung auf die Kompetenzniveaus: Lesekompetenz insgesamt

Im Folgenden wird berichtet, wie sich die Testleistungen der Schülerinnen und Schüler insgesamt am Ende der neunten Klasse auf die Kompetenzniveaus A bis C verteilen. Für eine Beurteilung der Niveaus ist dabei aus curricularer Sicht zu beachten, dass mit dem Eingangsniveau A die Mindestanforderungen des Hauptschulbildungsgangs markiert sind. Lesefähigkeiten unter diesem Niveau werden von den Leseverstehenstests in DESI nicht mit eigenen Test-Items erfasst.

In der Testentwicklung in DESI wurde der angestrebte curriculare Mindeststandard ursprünglich durch das Kompetenzniveau A auf einer Skala von A bis D markiert (siehe oben). Dieses Eingangsniveau A wurde im Zuge der Testauswertung mit dem ursprünglichen Niveau B zu einem erweiterten Gesamtniveau A auf einer Skala von A bis C zusammengefasst (vgl. Abbildung 12.4). Dementsprechend sind in dem erweiterten Kompetenzniveau A sowohl die Anforderungen des curricularen Mindeststandards der Hauptschule als auch etwas darüber hinaus gehende Anforderungen integriert.

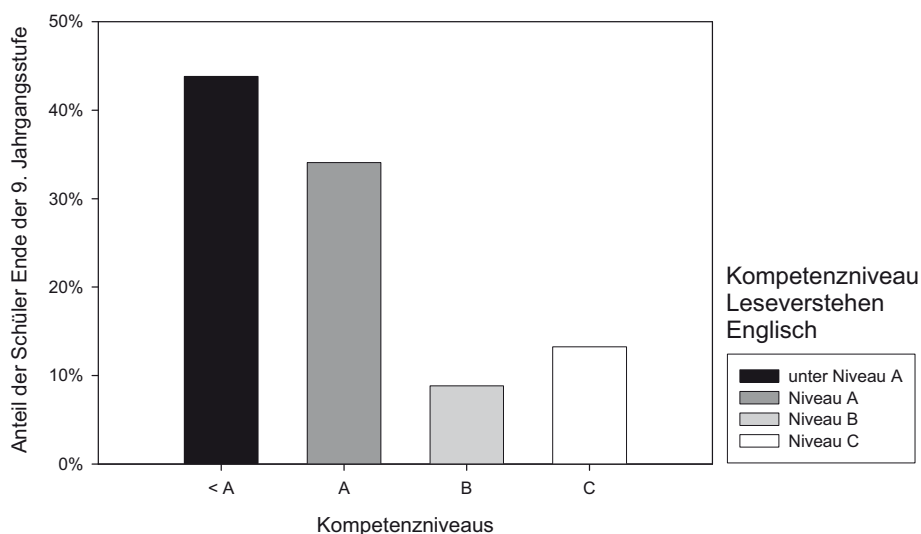


Abbildung 12.4: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Leseverstehen Englisch am Ende der neunten Jahrgangsstufe.

Wie Abbildung 12.4 zeigt, sind etwa 44% der Schülerinnen und Schüler nicht in der Lage, die Aufgaben des curricular angestrebten ersten Kompetenzniveaus hinreichend sicher zu lösen. Sie verfügen zwar über grundlegende Lesefähigkeiten (vgl. entsprechende Beschreibungen des Niveaus A1 des Gemeinsamen Europäischen Referenzrahmens für Sprachen; Europarat 2001); diese sind jedoch sprachlich zu begrenzt, um aus Texten im Umfang von bis zu einer Seite (270-400 Wörter), wie sie in DESI verwendet wurden, Einzelinformationen entnehmen zu können.

Ein gutes Drittel (34%) der Schülerinnen und Schüler erreichen das Niveau A, und zwar das Eingangsniveau A (8%) und das etwas erweiterte A+ (26%), und verfügen damit über eine grundlegende Kompetenz, die die Entnahme von Einzelinformationen aus sprachlich und thematisch auf die Erfahrungswelt der Schülerinnen und Schüler abgestimmten Texten erlaubt. Sie bewältigen in unterschiedlichem Maße curriculare Mindestanforderungen, und zwar bezogen auf das Ende des Hauptschulbildungsgangs (Niveau A und teilweise A+) sowie den mittleren Bildungsabschluss (Niveau A+ und B).

Mit 22% der Schülerinnen und Schüler ist auf den Niveaus B (9%) und C (13%) eine größere Gruppe in der Lage, Informationen beim Lesen zu verknüpfen, um Hauptaussagen zu verstehen und ein umfassendes mentales Modell eines entsprechenden Textes zu entwickeln (vgl. Kintsch 1998). Sie erreichen damit in unterschiedlichen Graden eine für die neunte Klasse schon deutlich weiter fortgeschrittene Lesekompetenz in der Fremdsprache.

Abschließend ist zu bemerken, dass möglicherweise eine begrenzte Anzahl von Schülerinnen und Schülern vor allem mit schwächerer Lesekompetenz Schwierigkeiten hatten, die leichten Aufgabenitems des Eingangsniveaus A in einem Text von etwa einer halben Seite und zusätzlichen Test-Items auch des Typs A+ zu erkennen und in der erwarteten Weise zu bearbeiten. Es ist daher nicht auszuschließen, dass die Ergebnisse bei Einsatz von Tests mit sprachlich einfachen Kurztexten von ca. fünf Zeilen und ausschließlicher Verwendung von Test-Items des Typs A (ohne A+) geringfügig besser ausgefallen wären.

13.4 Verteilung auf die Kompetenzniveaus: Unterschiede zwischen den Bildungsgängen

Ein Blick auf die Verteilung der Kompetenzniveaus in den verschiedenen Bildungsgängen zeigt, dass sich die Leistungen der Schülerinnen und Schüler im Lesen in der Fremdsprache Englisch sehr unterschiedlich auf die Kompetenzniveaus verteilen (vgl. Abbildung 12.5). Wie in anderen Englischtests in DESI stehen sich dabei einerseits die Kompetenzen in der Hauptschule und der Integrierten Gesamtschule und andererseits jene der Realschule und des Gymnasiums tendenziell gegenüber.

Bezogen auf den Mittelwert ihrer Leistungen beim Lesen liegen die Schülerinnen und Schüler der Hauptschule (430 Punkte) und der Integrierten Gesamtschule (444 Punkte) unterhalb von Kompetenzniveau A, das etwa bei 475 Punkten beginnt, wie aus Abbildung 12.5 zu entnehmen ist. Sie erreichen im Durchschnitt nicht die für dieses Kompetenzniveau notwendige Sicherheit im Lösen der Lesetestaufgaben.

Andererseits ist der Anteil der Schülerinnen und Schüler, die in diesen Bildungsgängen (Hauptschule und Integrierte Gesamtschule) die grundlegenden curricularen Anforderungen des Kompetenzniveaus A sowie die etwas darüber hinausgehenden Anforderungen des Niveaus A+ bewältigen, vergleichbar groß: An der Hauptschule sind dies insgesamt ca. 22% der Schülerinnen und Schüler, an der Integrierten Gesamtschule 26% (vgl. Abbildung 12.5). Ein kleiner Teil der

Schülerinnen und Schüler erreicht darüber hinaus die höheren Kompetenzniveaus B und C (2% an der Hauptschule bzw. 6% an der Gesamtschule), die deutlich über das curriculare Mindestniveau hinausgehen.

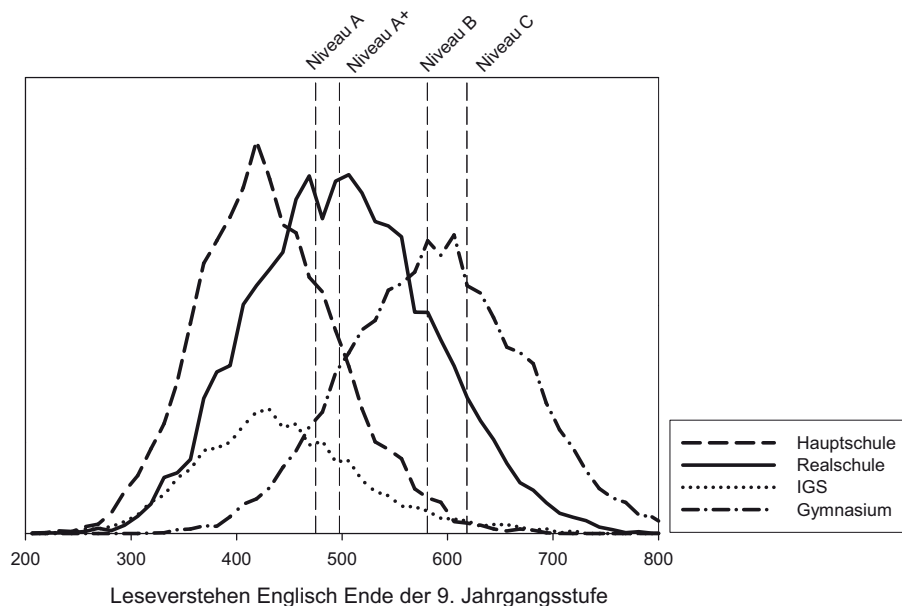


Abbildung 12.5 : Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Leseverstehen Englisch am Ende der neunten Jahrgangsstufe nach Bildungsgängen mit Binnendifferenzierung zwischen Kompetenzniveau A und A+.

Im Bildungsgang der Realschule erreichen die Schülerinnen und Schüler eine durchschnittliche Testleistung (501 Punkte), die dem Kompetenzniveau A/A+ entspricht (vgl. Abbildung 12.5). Knapp 40% der Schülerinnen und Schüler an Realschulen können die Aufgaben des ersten Kompetenzniveaus allerdings nicht ausreichend sicher lösen. Auf die Kompetenzniveaus B und C verteilen sich ca. 18% der Schülerinnen und Schüler (vgl. Abbildung 12.5).

Am Gymnasium haben die Schülerinnen und Schüler am Ende der neunten Klasse durchschnittlich eine Kompetenz entwickelt, die in der Tendenz Kompetenzniveau B zuzuordnen ist (589 Punkte), wie aus Abbildung 12.5 zu ersehen ist. Ein gutes Drittel der Gymnasiasten verteilt sich auf das oberste Kompetenzniveau C. Es ist davon auszugehen, dass diese Schülerinnen und Schüler in der Lage sind, implizite Informationen beim Lesen, wenn nötig, mit Hilfe von Weltwissen zu entschlüsseln und auch komplexe Informationen zu verstehen. Knapp 9% der Gymnasiasten sind dagegen der Gruppe der Schülerinnen und Schüler zuzuordnen, die Aufgaben des ersten Kompetenzniveaus nicht ausreichend sicher lösen können.

12.5 Diskussion

Die PISA-Ergebnisse zur ersprachlichen Lesekompetenz (Deutsches PISA-Konsortium 2001; Baumert u.a. 2002) haben offenbart, dass 20% der Fünfzehnjährigen zur Gruppe der schwachen und schwächsten Leser zählen (vgl. Artelt u.a. 2001). Die Ergebnisse zur Lesekompetenz Deutsch in DESI (vgl. Kapitel 5) zeigen ein etwas verändertes Gesamtbild; es wird jedoch auch hier deutlich, dass eine sehr große Anzahl von Schülerinnen und Schülern nur über eine sehr grundlegende und begrenzte Lesefähigkeit im Deutschen verfügen. Es ist vor diesem Hintergrund nicht überraschend, dass angesichts der zusätzlichen sprachlichen Anforderungen ein großer Anteil (44%) der deutschen Neuntklässler Schwierigkeiten beim Lesen in der Fremdsprache hat und die curricularen Mindestanforderungen nicht durchgängig bewältigen kann. Zur Erklärung dieses Befundes finden sich in der Lese- und Bilingualismusforschung unter anderem Hinweise, dass Lesestrategien der Erstsprache, sofern sie ausgebildet sind, nur dann in das fremdsprachliche Lesen transferiert werden können, wenn Leser ein bestimmtes sprachliches Schwellenniveau in der Erstsprache erreicht haben (vgl. *Threshold-Theorie*; Cummins 1979, 1991, 2000; Harley u.a. 1990); zugleich ist dafür jedoch auch in der Zweitsprache ein ausgeprägtes sprachliches Niveau unabdingbar (Baker 2006, S. 170 ff.; Urquart/Weir 1998; Alderson/Urquart 1984), und zwar zur Bewältigung von Sprache in kontextreduzierter Kommunikation (CALP), die vor allem für schulische Kontexte erforderlich ist (siehe die Unterscheidung zwischen mehr umgangssprachlicher (BICS) und kognitiv herausfordernder Kommunikation (CALP) bei Cummins/Swain 1986; vgl. auch Baker 2006, S. 174 ff.). Obgleich dieses Ergebnis auf Forschung mit einem anderen Zweitsprachenhintergrund (z.B. Immersionskontexte in Kanada) basiert, sind die Gemeinsamkeiten groß genug, um die Ergebnisse zur Erklärung auch hier heranzuziehen.

Es kann davon ausgegangen werden, dass viele der Schülerinnen und Schüler unterhalb des curricularen Niveaus über basale Lesefähigkeiten verfügen. Indirekt lassen die im unteren Kompetenzbereich klar differenzierenden Ergebnisse des C-Tests in DESI (vgl. Kapitel 14) darauf schließen, dass auch für die Lesekompetenz im Englischen mit einer gewissen, wenn auch sehr eingeschränkten Lesefähigkeit (vgl. GERS Niveau A1) zu rechnen ist. Da das DESI-Lesetestkonstrukt auf das curriculare Niveau der Hauptschule als Ausgangsniveau angelegt ist (vgl. Deutsche Bildungsstandards mit der Festlegung auf das GERS-Niveau A2), erfassen die entsprechenden Lesetests die Fähigkeiten darunter jedoch nicht, zumindest nicht inhaltlich beschreibbar auf der Lesekompetenzskala. Auf der anderen Seite wird sichtbar, dass die Lesekompetenz in der Fremdsprache Englisch in der neunten Klasse bei einer beachtlichen Anzahl von Schülerinnen und Schülern in verschiedenen Schularten schon sehr hoch entwickelt ist und bis in das für das Abitur angestrebte Niveau hineinreicht.

Literatur

- Alderson, J.C. (2000): *Assessing Reading*. Cambridge: Cambridge University Press.
- Alderson, J.C./Urquhart, A.H. (1984): *Reading in a foreign language*. New York: Longman.
- Artelt, C./Stanat, P./Schneider, W./Schiefele, U. (2001): Lesekompetenz: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske u. Buderich, S. 69-137.
- Baker, C. (*2006): *Foundations of Bilingual Education and Bilingualism*. Clevedon, Buffalo, Toronto, Sydney: Multilingual Matters.
- Baumert, J./Artelt, C./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Tillmann, K.-J./Weiß, M. (Hrsg.) (2002): PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich. Leske & Budrich: Opladen.
- Beck, B./Klieme, E. (Hrsg.) (2007): *Sprachliche Kompetenzen. Konzepte und Messung*. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Council of Europe (2001): *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press: Council of Europe.
- Cummins, J. (1979): Linguistic interdependence and the educational development of bilingual children. In: *Review of Educational Research* 49, S. 222-251.
- Cummins, J. (1991): Interdependence of first- and second-language proficiency in bilingual children. In: Bialystok, E. (Hrsg.): *Language processing in bilingual children*. Cambridge: Cambridge University Press, S. 70-89.
- Cummins, J. (2000): *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.
- Cummins, J./Swain, M. (1986): *Bilingualism in education*. New York: Longman.
- Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske u. Buderich.
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Grotjahn, R. (2000): Determinanten der Schwierigkeit von Leseverstehensaufgaben: Theoretische Grundlagen und Konsequenzen für die Entwicklung des TESTDAF. In: Bolton, S. (Hrsg.): TESTDAF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar. München: Goethe-Institut, S. 7-56.
- Harley, B./Allen, P./Cummins, J./Swain, M. (Hrsg.) (1990): *The development of second language proficiency*. Cambridge: Cambridge University Press.
- Kintsch, W. (1998): *Comprehension*. Cambridge: Cambridge University Press.
- Nold, G./Rossa, H. (2006): Anforderungsprofile von Aufgaben: Task-based assessment and task-based language learning. In: Hosenfeld, I./Schrader, F.-W. (Hrsg.): *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven*. Münster: Waxmann, S. 65-86.
- Nold, G./Rossa, H. (2007): Leseverstehen. In: Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen. Konzepte und Messung*. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 193-207.
- Nold, G./Rossa, H./Hartig, J. (in Druck): Proficiency scaling in DESI listening and reading EFL tests. Task characteristics, item-difficulty and cut-off points. Cambridge: *Studies in Language Testing*, Volume 27.
- Nold, G./Willenberg, H. (2006): Lesefähigkeit. In: Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen. Konzepte und Messung*. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 19-37.
- Smith, F. (1971): *Understanding reading: A psycholinguistic analysis of reading and learning to read*. New York: Holt, Rinehart, & Winston.
- Urquhart, A.H./Weir, C.J. (1998): *Reading in a Second Language: Process, Product and Practice*. New York: Longman.

Claudia Harsch / Konrad Schröder / Astrid Neumann

13 Schreiben Englisch

13.1 Testkonstrukt

Das DESI-Modul *semikreatives Schreiben* dient der Erfassung der Fähigkeit des *task*-orientierten Schreibens in der Fremdsprache (vgl. Harsch/Lehmann/Neumann/Schröder 2007). Dabei sollen den Schülerinnen und Schülern bei gleichzeitiger Lenkung möglichst große kommunikative Freiräume geboten werden. Das Testkonstrukt *Schreiben Englisch* ist in Analogie zum Konstrukt der *Textproduktion Deutsch* dreifach verortet in der linguistischen Text- und Schreibprozessforschung sowie in der didaktischen Forschung zur Schreibentwicklung, in der Analyse der relevanten Curricula aller Schularten und Bundesländer und in der empirischen Forschung zur Aufsatzbewertung. Das **fremdsprachliche Schreibkonstrukt in DESI** kann mit den Worten Camps definiert werden:

„Writing [is] a rich, multifaceted, meaning-making activity that occurs over time and in a social context, an activity that varies with purpose, situation, and audience and is improved by reflection on the written product and on the strategies used in creating it” (Camps 1996, S. 135).

Texte werden in diesem Konstrukt verstanden als funktionaler Versuch, dem Adressaten eine bestimmte Absicht zu übermitteln: „Writing may be said to represent an attempt to communicate with the reader.“ (Grabe/Kaplan 1996, S. 41)

Im Sinne des Performanz-Kompetenz-Modells sind die im Test gezeigten Leistungen als Indikatoren der Schreibkompetenz zu verstehen (zur Beschreibung des angesetzten Kompetenzmodells und der Ableitung der Bewertungskriterien vgl. auch Harsch u.a. 2007). Die Leistung resultiert aus den Fähigkeiten der Schülerinnen und Schüler und den Anforderungen der Aufgabenstellung. Deshalb werden im Folgenden zunächst die Testaufgaben vorgestellt und charakterisiert.

13.2 Testaufgaben

Die Curriculaanalysen ergaben, dass verständliches Schreiben in sinnvollen kommunikativen Zusammenhängen im Zentrum des Englischunterrichts in der neunten Jahrgangsstufe steht (vgl. Harsch u.a. 2007). Die Textsorten *persönlicher Brief* und (*Schülerzeitungs*)-*Bericht* können bildungsgangübergreifend eingesetzt werden. Als konsensfähige Themen erwiesen sich die Bereiche *Reise, Abenteuer und Erlebnisse, Beschreibung einer Person* und *Probleme Jugendlicher*. Folgende vier Aufgaben kamen zum Einsatz:

- 1: Bericht über eine Klassenfahrt ins englischsprachige Ausland,
- 2: Bericht über das Leben einer Person,
- 3: Brief an einen Jugendlichen, der Rat sucht,
- 4: Brief an die Großmutter über einen Londonbesuch.

In DESI wurde die fremdsprachliche Schreibkompetenz aus testökonomischen Gründen nur am Ende der neunten Jahrgangsstufe erfasst. Dabei bearbeitete jede Schülerin und jeder Schüler zwei Aufgabenstellungen: entweder den Bericht zur Klassenfahrt und den Brief an den Jugendlichen oder den Bericht zur Person und den Brief an die Großmutter.

13.3 Kodierung

Jeder Schülertext wurde im Doppel-Blind-Verfahren von je zwei Beurteilern bewertet. Die Beurteiler, mehrheitlich fortgeschrittene Lehramtsstudierende der Anglistik, durchliefen eine gründliche Schulung, in der sie mit Testkonstrukt, Messkonzept, Bewertungskriterien und Rating-Prozessen vertraut gemacht wurden. Die Texte wurden mittels Rating-Verfahren auf sechs Niveaus eingeschätzt (zur Ableitung und Definition der Bewertungskriterien und zur Beschreibung der *Rating Scales* und deren Niveaus vgl. Harsch u.a. 2007).

Insgesamt lagen 19490 Texte zur Auswertung vor. Die Güte der Bewertung lässt sich wie folgt beschreiben: Bei allen Bewertungskriterien zeigt sich, dass die Bewertungen in über 50% aller Fälle absolut übereinstimmen; zu etwas mehr als 40% unterscheiden sich die Ratings um eine Stufe. Damit können im Mittel mehr als 90% der Fälle mit hoher Übereinstimmung erfasst werden. Die hochsignifikanten Korrelationen zwischen Erst- und Zweitbewertung bewegen sich bei den taskspezifischen und sprachlichen Kriterien im Mittel im Bereich bei $r = 0.70$, beim Globalurteil im Mittel bei $r = 0.85$. Die Reliabilität der Bewertung liegt damit im akzeptablen Bereich (vgl. Harsch 2006).

13.4 Skalenbildung

Die Analysen ergaben für die fremdsprachliche Schreibkompetenz, dass es sinnvoll ist, einen Globalfaktor Schreiben über alle Aufgabenstellungen und Bewertungskriterien hinweg anzunehmen (vgl. Kapitel 8 sowie Neumann 2006a, 2006b). Faktorenanalysen bestätigten einen extrem engen statistischen Zusammenhang aller bewerteten Teilkriterien mit dem Globalurteil und lassen Rückmeldung in Form einer Globalskala *Schreiben* zu. Der den Auswertungen zugrunde liegende Schülerfähigkeitswert beruht dabei auf den vier Bewertungen des Gesamteindrucks der jeweils zwei von der Schülerin/dem Schüler bearbeiteten Aufgaben.

Den Beurteiler-spezifischen Effekten (Milde-/Strenge-Tendenzen) und den unterschiedlichen Anforderungen der Aufgaben wurde durch die zusätzliche Berücksichtigung dieser Aspekte bei der Skalierung mittels *partial-credit*-Modellen Rechnung getragen, um nicht Lernende zu benachteiligen, die von strengeren

Beurteilern bewertet wurden oder relativ betrachtet schwierigere Aufgaben zu bearbeiten hatten. Tabelle 13.1 zeigt die relativen aufgabenspezifischen Schwierigkeiten.

Tabelle 13.1: Aufgabenspezifische Schwierigkeiten für die Globalskala Schreiben Englisch in Punkten auf der DESI-Skala.

Aufgabenstellung	Schwierigkeit in Punkten DESI-Skala	nicht valide Texte in %
1: Bericht Klassenfahrt	566	9.90
2: Bericht Person	534	6.44
3: Brief Jugendlicher	605	12.64
4: Brief Großmutter	541	8.97

Es zeigt sich, dass den Schülerinnen und Schülern die Beschreibung einer Person respektive der Brief an die Großmutter, in dem über Erlebtes erzählt werden soll, vergleichbar leicht fällt. Das Verfassen eines Berichts über Abenteuer auf einer Klassenfahrt fällt deutlich schwerer; die höchsten Anforderungen scheint der Brief zu Problemen Jugendlicher zu stellen. Die Aufgabenschwierigkeit ist nicht systematisch an der Textsorte oder den Themen festzumachen. Es scheint, dass die Schwierigkeit u. a. vom Zugang zur Aufgabenstellung und der Bekanntheit der Thematik abhängt: Personenbeschreibungen (sei es nun die eigene Person oder fiktive Persönlichkeiten) werden im Unterricht geübt und sind wohlbekannt. Briefe an Freunde/Verwandte gehören ebenfalls zum unterrichtlichen Repertoire; zudem ist die Thematik eines London-Aufenthaltes hinreichend bekannt. Bei der Aufgabenstellung zum Bericht über die Klassenfahrt hingegen muss zunächst der Stimulus gedeutet werden, ehe die Schülerinnen und Schüler die besagte Klassenfahrt „erfinden“ müssen. Dies erfordert mehr Planung als die beiden erstgenannten Aufgaben. Die Schwierigkeit der Aufgabe, Probleme Jugendlicher zu verstehen und Ratschläge dazu zu geben, ist einerseits im englischsprachigen Stimulus zu verorten; andererseits erfordert die Aufgabenstellung Einfühlungsvermögen und die Entwicklung von Lösungen; dies stellt höhere Anforderungen an die Verarbeitungskapazitäten der Schülerinnen und Schüler und scheint ihnen vergleichsweise schwerer zu fallen als die erstgenannten Aufgabenstellungen.

Interessant ist, dass die Anteile der nicht validen Lernertexte mit der relativen Schwierigkeit der Aufgabe ansteigen. Sind bei der leichtesten Aufgabenstellung (2: Bericht Person) 6% der Texte nicht valide im Sinne des Testkonstrukts, so verdoppelt sich der Anteil bei der relativ schwersten Aufgabe auf 13%.

13.5 Kompetenzniveaus

Die globale Schreibkompetenzskala wurde über alle Aufgabenstellungen hinweg gebildet. Die Werte der Skala wurden auf die 500/100-Metrik transformiert. Aus der Skalierung ergaben sich die in Tabelle 13.2 dargestellten Niveaus.

Tabelle 13.2: Niveaus und Kompetenzbereiche der Globalskala Schreiben Englisch.

Niveau	Schwelle	Kompetenzbereich
Unter A	< 387.3	unter 385
A	387.3	385-460
B	463.2	460-555
C	555.2	555-675
D	674.0	675-770
E	769.4	über 770

Aus den Deskriptoren der taskspezifischen *Rating Scales* wurden die Deskriptoren für eine aufgabenübergreifende Kompetenzskala abgeleitet (zur Konstruktion der *Rating Scales* und der Kompetenzskala vgl. Harsch 2006). Dabei wurden die Niveaubeschreibungen mit relevanten Skalen aus dem *Gemeinsamen europäischen Referenzrahmen für Sprachen* (GERS)¹ abgeglichen. Dazu wurde der „Wortlaut der Deskriptoren“ (vgl. Europarat 2001) der GERS-Skalen mit den Beschreibungen der DESI-Niveaus verglichen. Es zeigten sich Berührungspunkte zwischen den DESI-Niveaus A bis D und den GERS-Niveaus A1 bis B2. Im obersten Bereich gab es teils Anknüpfungspunkte zu GERS-Niveau B2+, teils zu GERS-Niveau C1. Wo sich inhaltlich-qualitative Berührungspunkte zeigten, wurden die Niveaubeschreibungen an die Formulierungen des GERS angelehnt. Im DESI-Projekt kann jedoch eine empirische Anbindung, wie sie etwa das *Manual* (Council of Europe 2003) vorschlägt, nicht erfolgen. Deshalb handelt es sich bei dem hier erwähnten Abgleich lediglich um eine erste qualitative Einschätzung.

Ein nachträglicher Vergleich der Bildungsstandards der KMK zu Hauptschulabschluss (HS, angelehnt an GERS-Niveau A2) und Mittlerem Bildungsabschluss (MBA, angelehnt an GERS-Niveau B1) ergab folgende inhaltlich-qualitative Anknüpfungspunkte mit den DESI-Niveaus B und C:

DESI-Niveau B beinhaltet die beiden folgenden Beschreibungen aus den Bildungsstandards zum Hauptschulabschluss: „Die Schülerinnen und Schüler können:

- einfache, persönliche Briefe und E-Mails schreiben (A2),
- nach sprachlichen Vorgaben kurze einfache Texte (Berichte, Beschreibungen, Geschichten, Gedichte) verfassen (A2).“ (KMK 2004).

DESI-Niveau C umfasst die folgenden Formulierungen aus den Bildungsstandards für den Mittleren Bildungsabschluss: „Die Schülerinnen und Schüler können:

- in persönlichen Briefen Mitteilungen, einfache Informationen und Gedanken darlegen (B1),
- einfache standardisierte Briefe und E-Mails adressatengerecht formulieren, z. B. Anfragen, Bewerbungen (B1),

1 Relevant haben sich in diesem Zusammenhang folgende GERS-Skalen erwiesen: Schriftliche Produktion allgemein (S. 67), Kreatives Schreiben (S. 67f), Berichte und Aufsätze schreiben (S. 68), Schreiben (aus Raster zur Selbstbeurteilung, S. 36), Schriftliche Interaktion (S. 86), Korrespondenz (S. 86), Lexik (S. 112f), Grammatik (S. 114), Orthographie (S. 118), Kohäsionsmittel (S. 125), Spektrum allgemein (S. 110).

- unkomplizierte, detaillierte Texte zu einer Reihe verschiedener Themen aus ihren Interessengebieten verfassen, z. B. Erfahrungsberichte, Geschichten, Beschreibungen (B1),
- kurze einfache Aufsätze zu Themen von allgemeinem Interesse schreiben (B1).“ (KMK 2003).

Auch bei diesem Vergleich handelt es sich um eine inhaltlich-qualitative Einschätzung; es ist zu bedenken, dass die DESI-Testmodule keine Operationalisierung der Bildungsstandards darstellen, da letztere bei Testentwicklung noch nicht vorlagen. In Tabelle 13.3 stellen wir die Niveaus A, C und E der sechsstufigen DESI-Kompetenzskala exemplarisch vor; Formulierungen, die sich auf die Deskriptoren des GERS stützen, sind kursiv gedruckt. Das Niveau unter A wird nicht eigens beschrieben, da auf diesem Niveau kein Brief/Bericht verfasst werden kann.

13.6 Ergebnisse

Verteilung auf die Kompetenzniveaus am Ende der neunten Jahrgangsstufe

Nach der Darstellung des Testkonzepts und der Kompetenzniveaus stellen wir nun die Ergebnisse des Tests vor. Die Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus am Ende der neunten Jahrgangsstufe geht aus Abbildung 13.1 hervor.

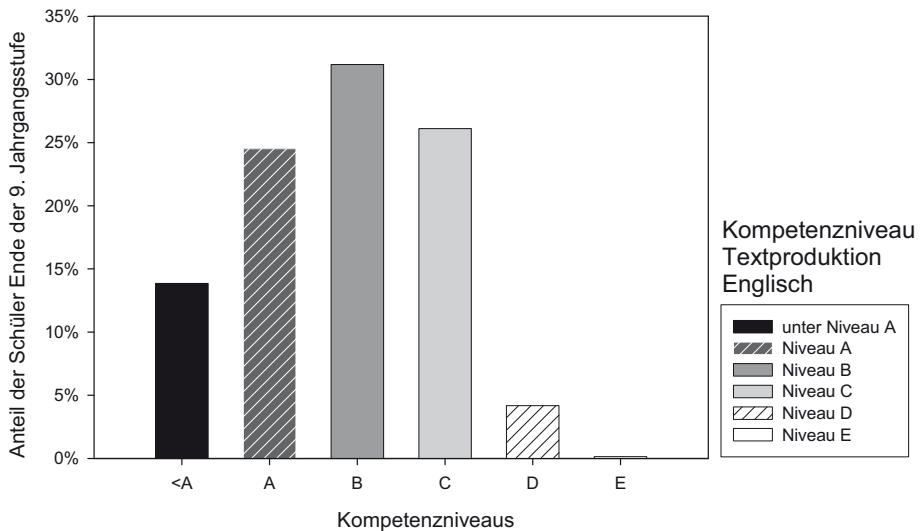


Abbildung 13.1: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Schreiben Englisch am Ende der neunten Jahrgangsstufe.

Tabelle 13.3: Beschreibung der Kompetenzniveaus A, C und E in Schreiben Englisch mit Aufgabenbeispielen.

Kompetenz-niveau	Beschreibung
E (über 770 Punkte)	<p>Kann <i>klare, detaillierte, gut strukturierte und ausführliche</i> Texte zu <i>komplexen Themen</i> verfassen wie beispielsweise eine Biographie schreiben, zwischenmenschliche Problemstellungen und deren Ursachen und Auswege erläutern oder Erlebnisse bei einem Auslandsaufenthalt beschreiben. Kann die eigene Ansicht darstellen. Kann <i>dabei die entscheidenden Punkte hervorheben, Standpunkte ausführlich darstellen und durch Unterpunkte oder geeignete Beispiele oder Begründungen stützen und den Text durch einen angemessenen Schluss abrunden</i>. Kann <i>Einstellungen und Gefühle</i> adäquat versprachlichen. Kann unterschiedliche Standpunkte einnehmen und verschiedene Perspektiven deutlich machen.</p> <p>Kann Berichte oder Briefe durchgängig entsprechend der <i>geltenden Konventionen</i> schreiben, wobei der <i>logische, stringente und konsistente Aufbau</i> das Verständnis erleichtert.</p> <p>Verfügt über einen <i>großen</i> Umfang sprachlicher Mittel und kann sie angemessen und variiert verwenden; verfügt über <i>Kollokationen</i> und <i>idiomatische</i> Wendungen; Fehler sind <i>selten</i>, können rückblickend korrigiert werden und sind <i>nicht kommunikationsbelastend</i>. Kann durchgehend <i>flüssige</i> Texte in <i>lesergerechtem, überzeugendem, persönlichem und natürlichem Stil</i> verfassen, welche narrative, <i>emphatische, humoristische</i> oder spannende Qualität aufweisen. Kann die angestrebte kommunikative <i>Wirkung</i> umfassend erzielen.</p>
C (555-675 Punkte)	<p>Kann <i>unkomplizierte, zusammenhängende</i> Texte zu <i>einer Reihe verschiedener Themen</i> aus der persönlichen Lebensumwelt, über <i>reale wie fiktive</i> Personen und <i>Ereignisse</i>, zwischenmenschliche Konstellationen und Probleme, <i>persönliche Erfahrungen</i> und einfache Sachverhalte schreiben. Interessenschwerpunkte können sichtbar werden. Kann zwischenmenschliche Beziehungen, <i>Gefühle und Reaktionen darauf in unkomplizierter Weise</i> beschreiben. Kann einfache, aber begründete Ratschläge geben und Lösungen zu alltäglichen Problemstellungen anbieten.</p> <p>Kann in einem <i>üblichen Standardformat</i> unkomplizierte <i>Berichte</i> oder Briefe schreiben, die eine logische Struktur aufweisen.</p> <p>Kann sprachliche Mittel bis zu einem <i>gewissen Grad</i> angemessen und in <i>hinreichendem Umfang</i> einsetzen, sofern sie sich auf <i>vorhersehbare</i> Situationen beziehen. <i>Zeigt eine gute Beherrschung des Grundwortschatzes und der gängigen grammatischen Strukturen</i>. Gewisse narrative Grundqualitäten (wie etwa Textkohärenz) sind gegeben. <i>Fehler</i> können im Text vorhanden sein, sie <i>schränken das Textverständnis</i> jedoch nur <i>gelegentlich</i> ein.</p> <p>Kann die Botschaft <i>grundsätzlich</i> kommunikativ <i>wirksam</i> vermitteln, doch teils mit <i>Einschränkungen</i>.</p>
A (385-460 Punkte)	<p>Kann <i>einfache Wendungen und Sätze über sich selbst</i> und andere (auch <i>fiktive</i>) <i>Menschen</i> schreiben: wie sie zusammengehören, <i>wo und wie sie leben, was sie tun</i> oder was sie tun wollen, sollen oder können.</p> <p>Kann <i>kurze, einfachste, persönliche</i> Berichte oder Briefe schreiben, die Mängel im Formalen zeigen. Kann dabei die Gedanken assoziativ niederschreiben.</p> <p>Zeigt das Vorhandensein <i>begrenzter sprachlicher Mittel aus dem hochfrequentem Basisbereich</i>, wobei die Texte meist bruchstückhaft und fehlerhaft sind (lexikalische, grammatische, syntaktische, orthographische Fehler, starke muttersprachliche Interferenz), weshalb die gewünschte Botschaft nur ansatzweise und oft missverständlich vermittelt werden kann.</p>

Dabei zeigt sich, dass 62% der Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe auf oder über dem Niveau B verortet sind und somit in der Lage sind, einen persönlichen Brief oder einen Schülerzeitungsbericht mehr oder weniger kommunikativ wirksam zu verfassen. Über 30% dieser Lernenden schaffen es, einen kurzen und einfachen Text zu erstellen und darin in meist linearer Reihung persönlich Erlebtes oder Alltägliches darzustellen, wobei die kommunikative Wirkung in groben Zügen erreicht wird. Dies mag zur einfachen Alltagskommunikation im privaten Umfeld genügen, auf fremdsprachliches Schreiben jenseits dieser Bereiche sind diese Jugendlichen jedoch nicht vorbereitet. Gut ein Viertel der Schülerinnen und Schüler (26%) kann unkomplizierte Texte in hinreichend korrekter und meist angemessener Sprache verfassen und die kommunikative Botschaft grundsätzlich wirksam vermitteln, wie man es in der neunten Jahrgangsstufe auch erwarten kann. 4% der Schülerinnen und Schüler (Kompetenzniveau D) verfügen über ein solch ausgeprägtes Sprachvermögen, das sie befähigt, klare und detaillierte Texte zu verschiedenen Themen zu verfassen, wobei Formalia eingehalten werden und ein breit gefächertes und angemessenes Sprachvermögen gezeigt wird. 0.1% aller Lernenden der neunten Jahrgangsstufe erreichen das oberste Niveau, das sich durch angemessene, gut strukturierte, flüssige Texte zu komplexen Themen in lesergerechtem Stil auszeichnet. Allerdings muss vermerkt werden, dass dieser Ausprägungsgrad der Schreibfähigkeit für die neunte Jahrgangsstufe in der Regel nicht erwartet werden kann.

Aufmerksamkeit verdienen die 38% der Lernenden, die auf Kompetenzniveau A und darunter verortet sind. Knapp ein Viertel dieser Jugendlichen (24%), diejenigen des Niveau A, können zwar einfache Wendungen und Sätze zu bekannten Themen schreiben, doch eine kommunikativ wirksame Botschaft oder gar Adressatenorientierung ist nicht festzustellen. Die 14% der Schülerinnen und Schüler, die unter Niveau A liegen, bedürfen dringend der Förderung bei der Entwicklung ihrer Schreibfertigkeit, denn sie sind nicht in der Lage, einen fremdsprachlichen Text zu produzieren, der bewertet werden könnte: Entweder sind diese Texte sprachlich dergestalt, dass nicht verständlich wird, was zum Ausdruck gebracht werden sollte, oder sie sind so kurz, dass sie nicht bewertet werden können.

Verteilung nach Bildungsgängen

Nun berichten wir, wie sich die Leistungsverteilung nach Bildungsgängen zusammensetzt: Wie aus Abbildung 13.2 ersichtlich, ist die Verteilung innerhalb der Bildungsgänge erwartungsgemäß: Gymnasiasten erzielen im Durchschnitt 595 Punkte auf der Rasch-Skala und liegen damit auf dem Kompetenzniveau C; sie können unkomplizierte Berichte und persönliche Briefe zu einer Reihe von Themen erstellen und dabei sprachliche Mittel hinreichend korrekt einsetzen. Realschüler erreichen im Mittel 500 Punkte, welche dem Kompetenzniveau B entsprechen; auf diesem Niveau können kurze, lineare Texte und persönliche Briefe zu vertrauten Themen in einfacher Sprache verfasst werden. Lernende an der integrierten Gesamtschule erzie-

len im Durchschnitt 446 Punkte und Hauptschüler 414 Punkte; diese Werte entsprechen dem Kompetenzniveau A, auf welchem kurze, einfachste Texte mit begrenzten sprachlichen Mitteln verfasst werden können.

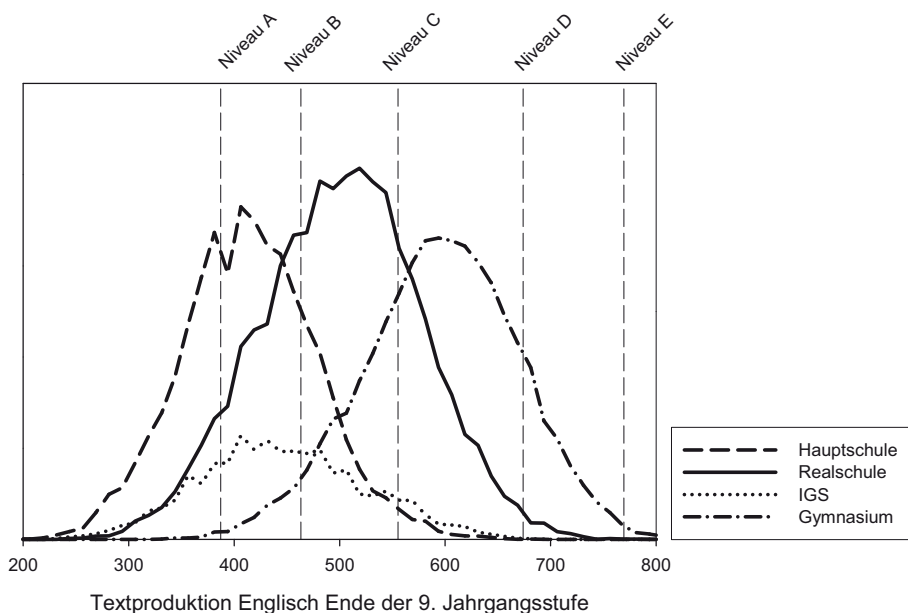


Abbildung 13.2: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Schreiben Englisch am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Am Gymnasium befinden sich 72% der Lernenden auf Kompetenzniveau C oder darüber, knapp ein Viertel (24%) sind auf Niveau B verortet und immerhin 4% der Lernenden liegen auf Niveau A oder darunter. Die Lernenden des Niveaus E sind allesamt am Gymnasium zu finden: 0.5% der Gymnasiasten können adressatengerechte, stilistisch angemessene Texte formulieren und dabei ein breites Spektrum sprachlicher Mittel korrekt verwenden.

An der Realschule ist die Mehrheit der Lernenden auf Niveau B zu finden (45%), 23% liegen auf Niveau C und noch 0.9% erreichen das Niveau D, auf welchem klare, detaillierte Texte in angemessener Sprache verfasst werden können. Allerdings liegen über 30% der Realschüler auf Niveau A oder darunter. Bei dieser Gruppe wären Fördermaßnahmen angeraten, um die schriftliche Kompetenz so weit zu entwickeln, dass am Ende der Schulzeit einfache Texte adressatenorientiert verfasst werden können.

Die Verteilung an Hauptschule und integrierter Gesamtschule verlangt indes nähere Betrachtung. In beiden Bildungsgängen kommen keine Lernenden über das Niveau C hinaus: Sind an der IGS noch 30% der Lernenden auf Niveau B und immerhin 10% auf Niveau C zu finden, so liegen an der Hauptschule in der Tat 20% der Schülerinnen und Schüler auf Niveau B und 2% auf Niveau C. Demnach kann davon ausgegangen werden, dass an der IGS etwa 40% und an der Hauptschule

ca. 20% der Lernenden am Ende ihrer Schulzeit in der Lage sein werden, einfache jedoch verständliche Texte zu persönlichen und bekannten Themen zu verfassen; damit erreichen sie vermutlich² die oben erwähnten Bildungsstandards für den Hauptschulabschluss. Hingegen sind 60% der Lernenden an der IGS und 78% der Hauptschüler nicht in der Lage, zusammenhängende schriftliche Texte in Englisch zu produzieren. Bei diesen Bildungsgängen müssten weit vor der neunten Klasse geeignete Maßnahmen getroffen werden, um die Jugendlichen wenigstens mit basalen Schreibkompetenzen in der Fremdsprache ins berufliche Leben zu entlassen.

Allerdings kommt es zu großen Überschneidungsbereichen, wie in Abbildung 13.2 dargestellt: Die stärksten Lernenden an Hauptschule und integrierter Gesamtschule etwa schneiden besser ab als die schwächsten Lernenden am Gymnasium. Die Leistungsstreuung ist in der integrierten Gesamtschule mit einer Standardabweichung von 78 ähnlich groß wie an der Realschule ($SD = 76$)³. Das Gymnasium zeigt mit einer Standardabweichung von 73 in etwa dieselbe Größenordnung, wohingegen sich an der Hauptschule ($SD = 64$) die geringste Streuung zeigt.

13.7 Implikationen für den fremdsprachlichen Schreibunterricht

Die oben dargestellten Befunde zur fremdsprachlichen Schreibkompetenz zeigen, dass es zu großen Unterschieden hinsichtlich der fremdsprachlichen Leistungen zwischen den Bildungsgängen in Deutschland kommt. Der Fremdsprachenunterricht am Gymnasium scheint zu hinlänglichen schriftlichen Kompetenzen zu führen, da schon am Ende der neunten Jahrgangsstufe über zwei Drittel der Lernenden in der Lage sind, zusammenhängende Texte zu produzieren, die sich in überwiegend angemessener Sprache auf die Adressaten beziehen. An der Realschule scheint es Bedarf zu geben, den Schreibunterricht zu optimieren, denn ein Jahr vor dem Schulabschluss sind immerhin 30% der Lernenden nicht in der Lage, verständliche und wirksame Texte zu erstellen. Zumindest wird hier jedoch bei der Mehrheit der Jugendlichen grundlegende Schreibkompetenz entwickelt, so dass sie einfache, zusammenhängende Texte schreiben können. An der integrierten Gesamtschule wird dies bei einem Drittel der Lernenden erreicht, wohingegen die Hauptschule gut 20% ihrer Schülerschaft mit dieser basalen Fähigkeit entlässt. Die Ergebnisse in den beiden letztgenannten Bildungsgängen dürften Aufmerksamkeit erregen – in diesem Bereich müssen kompensatorische unterrichtliche Maßnahmen ergriffen werden.

-
- 2 Eine Überprüfung der Bildungsstandards kann mit dem DESI-Testinstrumentarium nicht valide erfolgen, da die Standards wie erwähnt bei Erstellen der DESI-Instrumente noch nicht vorlagen. Aufgrund der oben ausgeführten Anknüpfungspunkte scheint es jedoch begründet, ein Erreichen der Standards zu vermuten, wenn das entsprechende DESI-Niveau erreicht wird. Diese Vermutung muss empirisch geprüft werden.
 - 3 Zur Veranschaulichung der Größe der Standardabweichung kann ein Vergleich mit der Ausdehnung des Kompetenzniveaus A helfen: Dieses umfasst 76 Punkte auf der Rasch-Skala.

Literatur

- Camp, R. (1996): New Views on Measurement and New Models for Writing Assessment. In: White, E./Lutz, W./Kamusiiki, S. (Hrsg.): *Assessment of writing: Politics, Policies, Practices*. New York: MLAA, S. 135-147.
- Council of Europe (2003): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). Manual. Preliminary Pilot Version*. Strasbourg: DGIV/EDU/LANG 5 rev. 1. 10. URL: http://www.coe.int/T/E/Cultural_Cooperation/education/Languages/Language_Policy/Manual (Zugriff am 13.9.2005).
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Grabe, W./Kaplan, R. B. (1996): *Theory and Practice of Writing: An Applied Linguistics Perspective*. New York: Longman.
- Harsch, C. (2006): *Der Gemeinsame europäische Referenzrahmen: Leistung und Grenzen. Die Bedeutung des Referenzrahmens im Kontext der Beurteilung von Sprachvermögen am Beispiel des semikreativen Schreibens im DESI-Projekt*. Universität Augsburg: Inaugural-Dissertation (Online-Publikation). <http://www.opus-bayern.de/uni-augsburg/volltexte/2006/368/>
- Harsch, C./Lehmann, R. H./Neumann, A./Schröder, K. (2007): *Übergreifende Konzeptualisierung sprachlicher Kompetenzen: Schreibfähigkeit*. In: Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)*. Weinheim: Beltz, S. 42-62.
- Kultusministerkonferenz (2003): *Vereinbarung über Bildungsstandards für die erste Fremdsprache für den Mittleren Schulabschluss*. URL: http://www.kmk.org/schul/Bildungsstandards/1.Fremdsprache_MSA_BS_04-12-2003.pdf (Zugriff am 5.11.2005).
- Kultusministerkonferenz (2004): *Vereinbarung über Bildungsstandards für die erste Fremdsprache für den Hauptschulabschluss*. URL: http://www.kmk.org/schul/Bildungsstandards/Hauptschule_Erste%20Fremdsprache_BS_307KMK.pdf (Zugriff am 5.11.2005).
- Neumann, A. (2006a): *Stabilität von Raterurteilen über die Zeit – Anpassung an vorhandene Schülerleistungen. Auswertung zweier Replikationsstudien zu den Urteilen in „DESI-Textproduktion“*. In: *Empirische Pädagogik* 20, H. 3, S. 286-296.
- Neumann, A. (2006b): *Dimensionen der Schreibfähigkeit. Differenzierte Analysen der Texte aus DESI und LAU11/ULME1*. Universität Hamburg: Inauguraldissertation.

Claudia Harsch / Konrad Schröder

14 Textrekonstruktion Englisch

14.1 Testkonstrukt

Zur Erfassung der Textrekonstruktion wird ein C-Test-Modul eingesetzt. Ein C-Test besteht aus kurzen, in sich geschlossenen Texten, in denen zweite Worthälften nach bestimmten Prinzipien getilgt werden. Die Erschließung dieser Lücken erfordert antizipatorische Sprachverarbeitung und die Nutzung von sprachlichen Redundanzen (vgl. Coleman 1996): Je kompetenter die Schülerinnen und Schüler in der Fremdsprache sind, desto mehr und umso anspruchsvollere Lücken können sie schließen.

Im DESI-Projekt kam das Modul C-Test zu beiden Testzeitpunkten zum Einsatz, um Aussagen über Veränderungen im Lauf der neunten Klasse treffen zu können. Dabei wurden von den Lernenden jeweils vier Kurztexte bearbeitet, wobei darauf geachtet wurde, dass keine Schülerin/kein Schüler am Ende der neunten Jahrgangsstufe dieselben Texte bearbeitete.

Zum Testkonstrukt, zur Aufgabenbeschreibung und zur Beschreibung der Testanforderungen und des Könnens auf den einzelnen Kompetenzniveaus darf auf Harsch/Schröder (2007) verwiesen werden.

14.2 Kompetenzniveaus

Die Leistung im Bereich der Textrekonstruktion wird anhand von fünf Kompetenzniveaus charakterisiert, die in Tabelle 14.1 beschrieben und anhand von Aufgabenbeispielen veranschaulicht werden.

Ein Vergleich der DESI-Niveaus mit den Niveaus des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GERS) ergab inhaltlich-qualitative Anknüpfungspunkte¹: Dabei wurde der „Wortlaut der Deskriptoren“ (vgl. Europarat 2001) der GERS-Skalen mit den Beschreibungen der DESI-Niveaus verglichen. Die DESI-Niveaus A bis E zeigen Berührungspunkte mit den GERS-Niveaus A1 bis C1, doch im DESI-Projekt kann eine empirische Anbindung, wie sie etwa das Manual (Council of Europe 2003) vorschlägt, nicht erfolgen. Deshalb handelt es sich bei dem hier erwähnten Abgleich lediglich um eine erste qualitative Einschätzung.

¹ Folgende Kategorien des GERS haben sich dabei als relevant erwiesen: Lesen (GERS, S. 74f), Rezeptionsstrategien (78), Texte verarbeiten (98), Lexik (112f), Grammatik (114), Orthographie (118), Kohäsionsmittel (125), Spektrum allgemein (110).

Tabelle 14.1: Beschreibung der Kompetenzniveaus in Textrekonstruktion Englisch mit Aufgabenbeispielen.

Kompetenz-niveau	Beschreibung	Beispiele
E (ab 689 Punkten)	Verarbeitung praktisch aller Texte: Integrativer Einsatz aller Wissensbestände; Sinnerschließungs- und Lesetechniken gut entwickelt; weitgehend automatisierte Sprachverarbeitung. Beherrschung umfangreicher sprachlicher Mittel.	Text "Women": The idea (...) began ea_____ in this century (...) Das Item <i>early</i> ist in dieser Struktur einem mittleren Sprachniveau zuzuordnen; die Semantik lässt sich über den Kontext erschließen, wobei ein mentales Modell der Textaussage dabei helfen kann; die Lücke kann kontextuell geschlossen werden. Das Item befindet sich in einem schweren Text.
D (611 bis 688 Punkte)	Verarbeitung auch komplexerer narrativer und Sachtexte: Schließung von Lücken unter Zuhilfenahme von Erschließungstechniken; Rekonstruktion semantischer Zusammenhänge; Sinnüberprüfung des rekonstruierten Textes. Großer Wortschatz zu den meisten allgemeinen Themen. Gute Beherrschung auch komplexerer grammatischer Phänomene.	Text "Big Ben": B_____ they had to do without [them] for two days (...). Die kontrastive Semantik des gesuchten Items <i>but</i> muss über den unmittelbaren Satz hinaus erschlossen werden, so dass Kontext wie Kontext genutzt werden müssen, um den Konnektor (textuelles Strukturierungsmittel) korrekt zu ergänzen. Das Item befindet sich in einem komplexeren Sachtext.
C (511 bis 610 Punkte)	Verarbeitung unkomplizierter Texte auch in weniger bekannten Themengebieten: Schließung von Lücken, die sich auf gängige sprachliche Phänomene beziehen, unter Zuhilfenahme von Ko-/Kontext und Weltwissen. Gute Beherrschung des Grundwortschatzes und der gängigen grammatischen Strukturen.	Text "Women": (...) even t_____ strictest chauvinist or macho has to accept the fact [that there are women in almost every trade and profession (...)]. Hier muss ein gängiges grammatisches Element, der Artikel <i>the</i> , nach einem emphatischen Lexem ergänzt werden, wobei es nur diese eine korrekte Alternative gibt. Das zu ergänzende Item ist dem sprachlichen Basisbereich zuzuordnen. Die Semantik ist unmittelbar zugänglich.
B (460 bis 510 Punkte)	Verarbeitung von kurzen, einfachen Texten*: Schließung von Lücken, die sich auf sprachliche Phänomene des curricularen Kernbereichs beziehen und deren Semantik unmittelbar zugänglich ist. Begrenzter, hochfrequenter Wortschatz der Alltagssprache sowie ein begrenztes, auf Routinesituationen bezogenes Spektrum an grammatischen Mitteln.	Text "First Rabbit": [She] waited patiently for me t_____ come (...). Hier muss in einer nicht hochfrequenten Struktur das Infinitiv-Partikel <i>to</i> ergänzt werden; das Item ist dem sprachlichen Basisbereich zuzuordnen; die Semantik ist eindeutig; das Item kann aus dem Kontext gelöst werden und befindet sich in einem leichten Text.
A (384 bis 459 Punkte)	Schließung von Lücken des sprachlichen Basisbereichs: Überwiegend vertraute Phänomene, hochfrequent oder parallel zur Erstsprache. Elementarer Vorrat von Wörtern und Basisstrukturen; meist auswendig gelerntes Repertoire, auf konkrete Situationen bezogen.	Text "First Rabbit": [The rabbit] sat un_____ a tree in our gar_____. Hier müssen dem Deutschen ähnliche Phänomene (Präposition/Nomen) ergänzt werden in einer häufig genutzten Struktur; beide lexikalischen Items sind hochfrequent; die Semantik ist unmittelbar zugänglich.

Anmerkung: *Textverarbeitung bezieht sich hier auf das Erschließen und Rekonstruieren der Kurztexte, die dem C-Test zugrunde liegen.

14.3 Ergebnisse

Verteilung auf die Kompetenzniveaus am Ende der neunten Jahrgangsstufe

Nach der Darstellung des Testkonzepts und der Kompetenzniveaus stellen wir nun die Leistungen der Schülerinnen und Schüler vor. Im Durchschnitt wurden am Ende der neunten Jahrgangsstufe 500 Punkte auf der Rasch-Skala erzielt mit einer Standardabweichung von 100 Punkten. Die Verteilung der Lernenden auf die Kompetenzniveaus am Ende der neunten Jahrgangsstufe geht aus Abbildung 14.1 hervor.

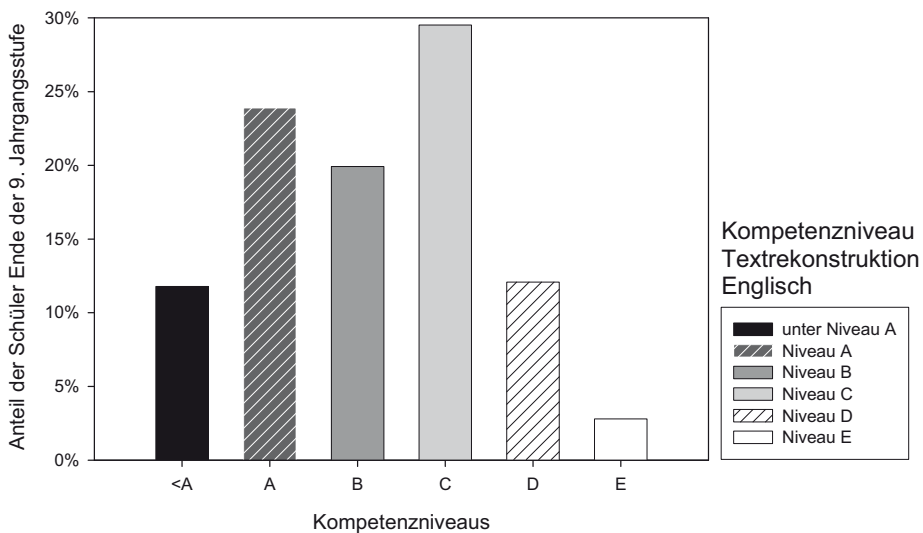


Abbildung 14.1: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Textrekonstruktion Englisch am Ende der neunten Jahrgangsstufe.

Dabei zeigt sich, dass 64% der Schülerinnen und Schüler der neunten Jahrgangsstufe oberhalb des Kompetenzniveaus A angesiedelt sind. Sie sind in der Lage, kurze Texte zu rezipieren, sie zu rekonstruieren und dabei fremdsprachliches Wissen und Texterschließungsstrategien in einem mehr oder weniger ausgeprägten Maß anzuwenden, so wie dies von einer neunten Jahrgangsstufe auch erwartet werden kann. Der fremdsprachliche Unterricht scheint zu Ergebnissen auf einem in seiner kommunikativen Funktionalität zumindest annehmbaren Niveau zu führen: Wie aus Abbildung 14.1 ersichtlich, erreichen 20% der Lernenden Niveau B und sind damit in der Lage, leichte, konkrete Texte, die sich auf alltägliche Themen beziehen, zu rekonstruieren. Ihr dabei eingesetztes sprachliches Können bezieht sich auf einen hochfrequenten Wortschatz der Alltagssprache und gebräuchliche grammatische Strukturen. 30% der Schülerinnen und Schüler, die Lernenden des Niveaus C, beherrschen über Grundwortschatz und Grundgrammatik hinaus auch weniger frequente Phänomene.

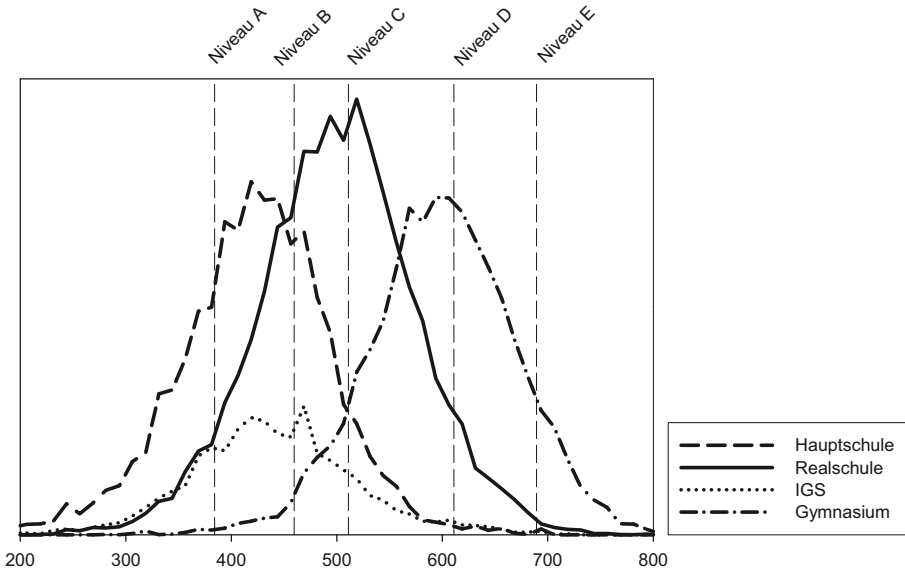
Sie können unkomplizierte Texte auch jenseits des eigenen Interessengebiets rezipieren und rekonstruieren.

Immerhin 12% der Schülerinnen und Schüler verfügen auf Kompetenzniveau D über solch ausgeprägte Lese- und Interpolationstechniken und können dabei ihre sprachlichen Wissensbestände so verschränkt und automatisiert einsetzen, dass sie auch komplexere Texte zu Themen jenseits der eigenen Lebenswelt erschließen und die rekonstruierten Texte auf ihren Sinn überprüfen können. Auf oberstem Niveau E, für das eine weitgehend automatisierte Sprachverarbeitung und eine effiziente Nutzung verschiedener Lese- und Sinnerschließungstechniken angesetzt werden, sind noch 3% der Lernenden der neunten Jahrgangsstufe zu finden: Sie verfügen über umfangreiche sprachliche Mittel einschließlich idiomatischer Wendungen und seltener grammatischer Phänomene, so dass sie alle Arten von Kurztexten erschließen und rekonstruieren können.

Vermehrte Aufmerksamkeit sollte den 36% der Lernenden zukommen, die auf Kompetenzniveau A oder darunter liegen. Die Lernenden auf Niveau A, das betrifft 24% der Schülerschaft, verfügen zumindest über basale Sprachverarbeitungskapazitäten und sind in der Lage, hochfrequente Phänomene in einem teils auswendig gelernten Repertoire anzuwenden. Die 12% der Lernenden jedoch, die noch unter diesem Niveau liegen, verfügen vermutlich nicht einmal über diese grundlegende Sprachbeherrschung.

Kompetenzverteilung nach Bildungsgängen

Die Leistungsverteilung im Sinne der Zuordnung der Schülerinnen und Schüler zu Kompetenzniveaus hängen mit dem Bildungsgang zusammen, wie aus Abbildung 14.2 ersichtlich wird. Am Gymnasium befinden sich 90% der Lernenden auf oder über dem Niveau C, davon immerhin 44% auf den beiden obersten Niveaus. Es kann demnach davon ausgegangen werden, dass die Gymnasiasten bereits am Ende der neunten Jahrgangsstufe über solide fremdsprachliche Kenntnisse verfügen und diese auch anwenden können. An der Realschule sind 44% der Lernenden auf oder über dem Niveau C verortet, 28% auf Niveau B und noch 28% der Lernenden darunter. Die Mehrheit der Realschüler verfügt demnach über Grundkenntnisse in der Fremdsprache, die sie zu alltäglicher Kommunikation bezogen auf ihnen bekannte Sachverhalte befähigen sollten. An Hauptschule und integrierter Gesamtschule zeichnet sich ein anderes Bild ab: Hier befinden sich 70% resp. 61% der Lernenden auf oder unterhalb des Niveaus A; 21% resp. 22% erreichen das Niveau B; das Niveau C und höher wird von 8% resp. 16% der Lernenden erreicht.



Textrekonstruktion Englisch Ende der 9. Jahrgangsstufe

Abbildung 14.2: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Textrekonstruktion Englisch am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Allerdings kommt es zu großen Überschneidungsbereichen: Die stärksten Lernenden an Hauptschule und integrierter Gesamtschule etwa schneiden besser ab als die schwächsten Lernenden am Gymnasium. Die Leistungsstreuung ist erwartungsgemäß in der integrierten Gesamtschule mit einer Standardabweichung von 79 am größten. Dies deutet auf ein besonders hohes Maß an Heterogenität der Schülerschaft hin. Hauptschule ($SD = 72$) und Realschule ($SD = 74$) weisen in etwa dieselbe Standardabweichung auf, wohingegen sich am Gymnasium ($SD = 70$) die geringste Streuung zeigt, vermutlich da sich hier die homogenste Schülerschaft befinden dürfte.

Kompetenzzuwachs in der neunten Jahrgangsstufe

Die Veränderungsmessung zu Beginn und gegen Ende der neunten Jahrgangsstufe ergab beim Modul C-Test die in Abbildung 14.3 ersichtlichen Leistungszuwächse. Es zeigt sich eine deutliche Veränderung der Verteilung auf die Kompetenzniveaus: Lagen zu Beginn der neunten Jahrgangsstufe noch 45% der Lernenden unter Kompetenzniveau B, so hat sich dieser Anteil zum Ende der neunten Jahrgangsstufe um 9% verringert. Auf Niveau B befinden sich in etwa gleich viele Schülerinnen und Schüler. Der Anteil derer, die auf dem Niveau C verortet werden, nimmt um 3% zu, der Anteil der Lernenden auf dem Niveau D sogar um 4%. Selbst auf dem ober-

sten Niveau E liegen am Ende der neunten Jahrgangsstufe 2% mehr Lernende als zu Beginn der neunten Jahrgangsstufe.

Demnach ist – über alle Bildungsgänge hinweg – ein Lernzuwachs an globalem Sprachvermögen in der neunten Jahrgangsstufe des deutschen Schulsystems zu verzeichnen. Um die Bedeutung dieses Zuwachses interpretieren zu können, werfen wir im Folgenden einen Blick auf die Leistungsverteilung und die Leistungszuwächse bezogen auf die verschiedenen Bildungsgänge.

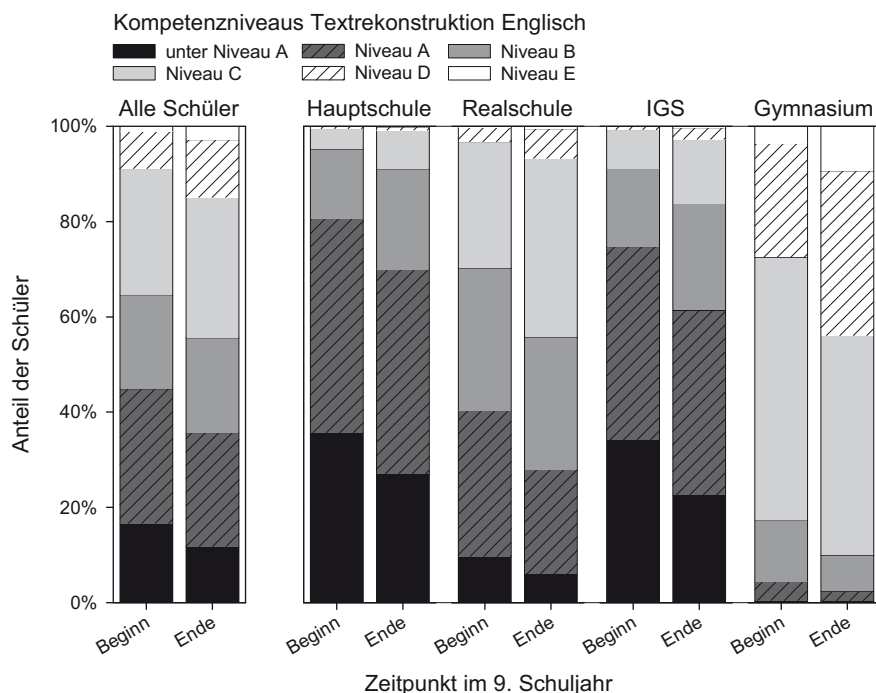


Abbildung 14.3: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Textrekonstruktion Englisch zu Beginn und am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Abbildung 14.3 zeigt einen deutlichen Unterschied der Zuwächse differenziert nach den einzelnen Bildungsgängen: Die Gymnasiasten erzielen zum zweiten Testzeitpunkt (am Ende der neunten Jahrgangsstufe) im Mittel 599 Punkte auf der Rasch-Skala, das sind 27 Punkte mehr als zum ersten Testzeitpunkt (zu Beginn der neunten Jahrgangsstufe); lag der Mittelwert zu Beginn der neunten Jahrgangsstufe noch im mittleren Bereich des Kompetenzniveaus C, so verschiebt er sich am Ende der neunten Jahrgangsstufe in den oberen Bereich des Niveaus C. An der Realschule werden am Ende der neunten Jahrgangsstufe durchschnittlich 500 Punkte erzielt, wohingegen der Mittelwert zu Beginn der neunten Jahrgangsstufe bei 475 Punkten lag; die Zunahme um durchschnittlich 25 Punkte entspricht einer Verschiebung des Mittelwerts vom unteren Ende des Kompetenzniveaus B hin zum oberen Ende dieses Niveaus. An der integrierten Gesamtschule erreichen die Schülerinnen und Schüler durchschnittlich 440 Punkte am Ende der neunten Jahrgangsstufe und erzielen da-

mit 28 Punkte mehr als zu Beginn der neunten Jahrgangsstufe. Die Lernenden lagen im Mittel zu Beginn der neunten Jahrgangsstufe auf Kompetenzniveau A, im Ende der neunten Jahrgangsstufe verschieben sie sich zum oberen Ende dieses Niveaus hin. An der IGS kommt es zur größten Leistungssteigerung im Vergleich zu den anderen Bildungsgängen. An der Hauptschule liegt der Mittelwert am Ende der neunten Jahrgangsstufe bei 421 Punkten; im Vergleich zum Beginn der neunten Jahrgangsstufe erzielen die Hauptschüler 15 Punkte mehr; sie liegen damit zu beiden Testzeitpunkten im Mittel auf Niveau A.

An der Hauptschule und der integrierten Gesamtschule gibt es die größte Verschiebung der Lernenden von den beiden untersten Niveaus auf die Niveaus B und C. Hier scheint es, der fremdsprachliche Unterricht vermag noch in der letzten Jahrgangsstufe vor dem Schulabschluss einen Teil der Lerner mit den schwächsten Leistungen so zu fördern, dass sie die Schule zumindest mit basalen fremdsprachlichen Kenntnissen verlassen, die sie befähigen, alltägliche Fremdsprachenanforderungen zu meistern. An der Realschule zeigt sich ebenfalls eine Verschiebung von den unteren Niveaus auf die oberen Niveaus; hier jedoch dürfte es sich um einen gleichmäßigen Lernzuwachs unter allen Lernenden handeln. Denn es zeigt sich ein Zuwachs über alle Niveaus hinweg, wobei es zur größten Zunahme im Mittelbereich kommt. Auch wenn zum Ende der neunten Jahrgangsstufe noch 28% der Realschüler unterhalb des Niveaus B liegen, so dürfte doch der Großteil der Realschüler am Ende der zehnten Klasse – vorausgesetzt, der Lernzuwachs verhält sich in der zehnten Jahrgangsstufe ähnlich dem in der neunten – auf Alltagssprachliche Kommunikation in der Fremdsprache vorbereitet sein. Am Gymnasium ist der Zuwachs im oberen Bereich am deutlichsten: Hier bewegen sich die Lernenden von den unteren und mittleren Niveaus hin zu den beiden obersten Niveaus.

14.4 Implikationen für den Fremdsprachenunterricht

Die Befunde zum globalen Sprachstand zeigen, dass es zu großen Unterschieden hinsichtlich der fremdsprachlichen Leistungen zwischen den Bildungsgängen in Deutschland kommt: Während die Ausbildung am Gymnasium zur Entwicklung einer insgesamt akzeptablen kommunikativen Kompetenz führt, zeigt sich in den anderen Bildungsgängen ein etwas anderes Bild: Die Realschule schafft es, die Mehrheit der Lernenden auf alltägliche Bedürfnisse der fremdsprachlichen Kommunikation vorzubereiten. Dennoch kann nicht übersehen werden, dass mehr als ein Viertel der Lernenden auf unterstem Niveau oder noch darunter angesiedelt sind. Diese Lernenden können vertraute, hochfrequente Phänomene verarbeiten und verfügen über elementarste sprachliche Kenntnisse; von einer Automatisierung der Sprachverarbeitung kann jedoch nicht die Rede sein. Hier sollte die Gruppe der Lernenden mit den schwächsten Leistungen gezielt gefördert werden.

Die Ergebnisse an Hauptschule und IGS dürften Aufmerksamkeit erregen: Hier führt die schulische Ausbildung bei nur einer Minderheit der Lernenden zur Herausbildung grundlegender kommunikativer Kompetenzen. Die Mehrheit erreicht

das unterste Kompetenzniveau nicht. Es muss davon ausgegangen werden, dass die Lernenden in diesen Bildungsgängen über ein so wenig ausgeprägtes Sprachvermögen verfügen, dass sie auch einfachste fremdsprachliche Anforderungen mehrheitlich nicht erfüllen können. Vermutlich muss hier zunächst einmal das Lernumfeld verändert werden, damit Lernmotivationen entstehen können, die zur Entwicklung kommunikativer Fertigkeiten zumindest in grundlegenden Bereichen des Alltags beitragen können.

Literatur

- Coleman, J.A. (1996): *Studying Languages: a survey of British and European students*. London: Centre for Information on Language Teaching and Research.
- Council of Europe (2003): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). Manual. Preliminary Pilot Version*. Strasbourg: DGIV/EDU/LANG 5 rev. 1. 10. URL: http://www.coe.int/T/E/Cultural_Cooperation/education/Languages/Language_Policy/Manual (Zugriff am 13. 9. 2005).
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Harsch, C./Schröder, K. (2007): *Kompetenzmodelle und Kompetenzniveaus im Bereich des Englischen: Textrekonstruktion C-Test*. In: Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen Konzepte und Messung DESI-Studie (Deutsch Englisch Schülerleistungen International)*. Weinheim: Beltz, S. 212-225.

Günter Nold / Henning Rossa

15 Sprachbewusstheit Englisch

Beim Lernen der Fremdsprache Englisch werden Schülerinnen und Schüler mit neuen kommunikativen und sprachlichen Regelungen konfrontiert. Gehen sie aufmerksam mit diesen Regelungen um, besteht für sie die Möglichkeit, besonders sprachbewusst zu werden. Die Sprachbewusstheit befähigt sie in vielen Fällen dazu, das gelernte Englisch kontrolliert anzuwenden, den Sprachgebrauch zu beurteilen und Verstöße gegen Regelungen zu erkennen sowie zu korrigieren.

Ausgehend von diesen Vorstellungen wird im DESI-Modul Sprachbewusstheit untersucht, in welchem Umfang die Schülerinnen und Schüler entsprechende Kompetenzen im Wesentlichen auf Grund ihrer unterrichtlichen Lernerfahrungen in sprachlichem Handeln und in Grammatik abrufen und einsetzen können. Eine Teilkompetenz ist dabei die Fähigkeit zur Sprachreflexion. Sie beruht in der Regel auf Wissen über Sprache (explizites/deklaratorisches/verbalisierbares Wissen), während die Teilkompetenz der sprachlichen Korrekturfähigkeit stärker vom Wissen des Sprachgebrauchs oder des Sprachgefühls (vorwiegend implizites und prozedurales Wissen) bestimmt ist (vgl. Karmiloff-Smith 1992; Cameron 1993; Ellis 1994a, 1994b, 2004; Schmidt 1995; Aplin 1997; Alderson/Clapham/Steel 1997; Han/Ellis, 1998; Hawkins 1999; Nold/Grimmig 2000; Paradis 2004; Eichler/Nold 2007). Da die Regelungen im Bereich des Sprachhandelns und der satzbezogenen Grammatik unterschiedlichen Bereichen der Sprache angehören und sich in der Fremdsprache deutlich in der Art, wie sie verfügbar sind, voneinander unterscheiden, wird im Modul Sprachbewusstheit von zwei getrennten Testkonstrukten ausgegangen. So werden Sprachbewusstheit Soziopragmatik und Sprachbewusstheit Grammatik jeweils mit eigenen Tests erfasst.

15.1 Testkonstrukt Sprachbewusstheit Soziopragmatik

Das Konstrukt im Bereich der Soziopragmatik beschreibt, welche Aspekte der Kompetenz in dem entsprechenden Test des DESI-Moduls erfasst werden. Im Mittelpunkt stehen einerseits sprachlich-soziale Fähigkeiten und andererseits Aspekte der Diskurskompetenz, insbesondere der Sinnkonstruktion in Texten. So soll erkannt und beurteilt werden, ob sprachliche Äußerungen stilistisch ausreichend bestimmten Situationen angepasst sind. Ferner geht es darum, Höflichkeitsanforderungen zu bewerten. Schließlich kommt es darauf an zu erkennen, welche Intentionen Sprecher in ausgewählten Bereichen ausdrücken und wie Sprechintentionen sich in den Zusammenhang fortlaufender sprachlicher Rede einordnen. Sprachbewusstheit als Kompetenz ist hier demnach darauf ausgerichtet, Facetten sprachlichen Handelns im Diskurs zu erkennen, einzuordnen und zu bewerten. Den Schülerinnen und Schülern wurde entsprechend dieser Konzeption von Sprachbewusstheit Gelegenheit zur nö-

tigen Reflexion gegeben. In realen Gesprächssituationen bleibt gewöhnlich wenig Zeit, über sprachliches Handeln zu reflektieren, vor allem wenn unmittelbar sprachlich gehandelt werden muss. In den Tests, die entsprechend diesem Testkonstrukt entwickelt wurden, erscheint dagegen die mündliche Rede in Schriftform, sodass die mit dem sprachlichen Handeln korrespondierenden Elemente der Sprachbewusstheit unter geringerem Zeitdruck abgerufen werden konnten.

Als verbindender Kontext wurde für die Mehrzahl der Testaufgaben die Situation des Schul- bzw. Schüleraustauschs gewählt, da dieser Kontext in der achten und neunten Klasse in vielen Schulen besonders relevant ist. Außerdem bietet dieser Kontext genügend Ansatzpunkte zu einem Transfer in berufsbezogene Situationen.

Kompetenzniveaus und Aufgabenbeispiele

Mit Hilfe der Testaufgaben ist es möglich, die Schülerinnen und Schüler bestimmten Kompetenzniveaus im Bereich der Sprachbewusstheit zuzuordnen. Diese Kompetenzniveaus basieren nämlich auf den Anforderungsprofilen der Testaufgaben; sie wurden verbunden mit der empirischen Aufgabenschwierigkeit analysiert und schließlich zur Markierung von drei Niveaus auf einer kontinuierlichen Skala genutzt. Vier der entwickelten Aufgabenmerkmale im Bereich der Sprachbewusstheit Soziopragmatik eigneten sich dazu, die beobachteten Unterschiede in den Aufgabenschwierigkeiten zu erklären. Zu diesen gehören Merkmale der Texte und der Textverarbeitung. Es zählen dazu im Einzelnen die Art des inhaltlichen Zusammenhangs und der unterschiedlich deutlichen textuellen Verknüpfung, ferner Fähigkeiten in der Erkenntnis und Unterscheidung von Sprachfunktionen sowie unterschiedliche Möglichkeiten, auf Vorwissen vor allem aus der Erstsprache zuzugreifen (vgl. Nold/Rossa 2007). Falls das erste Kompetenzniveau von einigen Schülerinnen und Schülern nicht erreicht wird, lassen sich demnach insgesamt vier Gruppen mit den drei Kompetenzniveaus unterscheiden (vgl. Tabelle 15.1).

Tabelle 15.1: Beschreibung der Kompetenzniveaus in Sprachbewusstheit Englisch Soziopragmatik.

Kompetenz-niveau	Beschreibung
C (ab 702 Punkte)	Ist fähig, sprachliches Handeln zu durchschauen, auch wenn der inhaltliche Zusammenhang stärker eigenes Erschließen erfordert. Zusätzlich kann es nötig sein, muttersprachlich bedingte Missverständnisse zu bewältigen und spezifisch fremdsprachliches Handlungswissen erschließend einzusetzen.
B (516-701 Punkte)	Ist fähig sprachliches Handeln zu durchschauen, auch wenn die Art des Handlungsablaufs und vorhandene textuelle Verknüpfungsmittel es notwendig machen, spezifisch fremdsprachliches Handlungswissens einzusetzen.
A (437-515 Punkte)	Ist fähig sprachliches Handeln zu durchschauen, wenn Handlungsablauf und textuelle Verknüpfungsmittel dem Lerner durch vergleichbare Texte vertraut sind.

In den folgenden Abbildungen 15.1 bis 15.3 werden Aufgabenbeispiele für die drei Kompetenzniveaus dargestellt. Eine detaillierte Darstellung findet sich bei Nold/Rossa (2007).

Begrüßungsszene in einer Firma 1

Frank: Good morning, Kate.
 Kate: Hello, Frank. It's good to see you. How are you?
 Frank: A ☐ Fantastic! I just got a new job.
 B ☐ I'm O.K.
 C ☐ I'm fine, thanks. And you?
 D ☐ Couldn't be better.
 Kate: I'm O.K. See you later.

Abbildung 15.1: Aufgabenbeispiel für Kompetenzniveau A in Sprachbewusstheit Englisch Soziopragmatik.

Bei dem Testitem in Abbildung 15.1 haben die Schülerinnen und Schüler eine der vier Antwortmöglichkeiten auszuwählen. Dabei kommt die Textkohärenz ebenso zum Zuge wie die Art der Verknüpfung und die Vertrautheit mit der Textsorte. Antwort C ist durch Paralleltexte so geläufig, der Textzusammenhang so vertraut, dass damit das erste Kompetenzniveau zutreffend erfasst wird.

Was sagst Du in folgender Situation?

Du erledigst deine Einkäufe auf einem Wochenmarkt in Edinburgh (Schottland). Nachdem du viel Obst eingekauft hast, packt der Obst- und Gemüsehändler noch einige Zwiebeln und Karotten oben drauf.
 Shopkeeper:
 Here's something else. And you can make a tasty soup with these.

You: Was würdest du antworten? Welche Antwort passt wohl am besten? Kreuze A, B, C oder D an. (Nur ein Kästchen ist richtig).

A ☐ Thank you. But I don't like onions.
 B ☐ Thanks anyway.
 C ☐ That's lovely, thanks a lot.
 D ☐ Is that usual in Scotland?

Abbildung 15.2: Aufgabenbeispiel für Kompetenzniveau B in Sprachbewusstheit Englisch Soziopragmatik.

Bei der Lösung der Aufgabe für das Niveau B (Antwort C) in Abbildung 15.2 werden die Schülerinnen und Schüler aufgefordert, sich in eine Einkaufssituation in Schottland zu versetzen. Vorerfahrungen aus ihnen vertrauten Kontexten und Paralleltexten können ihnen dabei durchaus helfen; sie können jedoch nicht ungeprüft vertraute Handlungsmuster übertragen, zumal situationsspezifische Höflichkeitserwartungen und auch sprachlich-idiomatische Aspekte zu berücksichtigen sind.

Before lunch

Du reagierst auf eine Äußerung deines Freundes/ deiner Freundin. Dabei kannst du dich **zustimmend** (z.B. „ich auch“) oder auch **nicht zustimmend** (z.B. „ich nicht“) ausdrücken. Wähle die am besten passenden Aussagen (A, B, C oder D) [...] aus. Es gibt **zwei Lösungen**. [...]

Friend: I'd like a ham and cheese sandwich.

You:

- A ☐ Never mind!
- B ☐ I didn't like a sandwich.
- C ☐ I'd like one as well.
- D ☐ I'd rather have a burger and fries.

Abbildung 15.3: Aufgabenbeispiel für Kompetenzniveau C in Sprachbewusstheit Englisch Soziopragmatik.

Die Aufgabe in Abbildung 15.3 entspricht Kompetenzniveau C, wenn Antwort D gewählt wird – die ebenfalls zutreffende Antwort C gehört dagegen zu Kompetenzniveau B (vgl. Nold/Rossa 2007). Der komplexere Textzusammenhang von Antwort D definiert hier den Schwellenwert im Übergang von Kompetenzniveau B zu C. Zugleich ist diese Aufgabe hinsichtlich der textuellen Verknüpfung und der Sprachfunktionen im oberen Bereich und macht sie daher für die Schülerinnen und Schüler zu einer schwierigen Aufgabe insgesamt.

Verteilung auf die Kompetenzniveaus in Sprachbewusstheit Soziopragmatik: Schülerleistungen insgesamt

Im Folgenden wird dargelegt, wie sich die Schülerinnen und Schüler auf die verschiedenen Niveaus im Bereich Sprachbewusstheit Soziopragmatik verteilen. Wie Abbildung 15.4 zeigt, verteilen sich knapp 30% der Schülerleistungen zu Beginn der neunten Jahrgangsstufe auf das Kompetenzniveau A; fast der gleiche Prozentsatz von Schülerleistungen liegt unter diesem Niveau. Die Schülerinnen und Schüler auf Niveau A können elementare Fähigkeiten einsetzen, wenn es darauf ankommt, mit soziopragmatischen Regelungen bewusst umzugehen. Wird dieses Niveau nicht erreicht, sind entsprechende elementare Fähigkeiten noch nicht konsistent ausgebildet. Ferner ist zu beobachten, dass etwas über 40% der Schülerleistungen Niveau B zurechnen sind, weitere zwei bis drei Prozent liegen auf Niveau C. Schon auf Niveau B gelingt es Schülerinnen und Schülern, spezifisches soziopragmatisches Wissen abzurufen und anzuwenden, um mit Fragen der Angemessenheit und Höflichkeit fremdsprachlichen Verhaltens umgehen zu können.

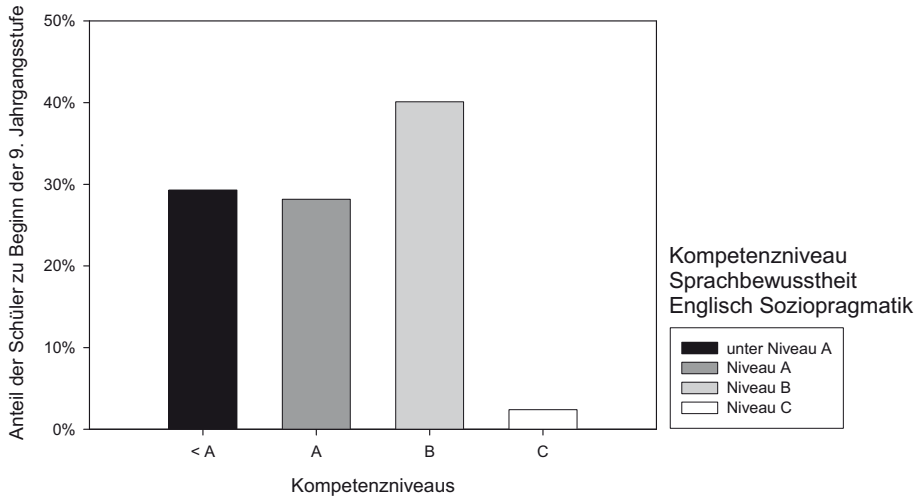
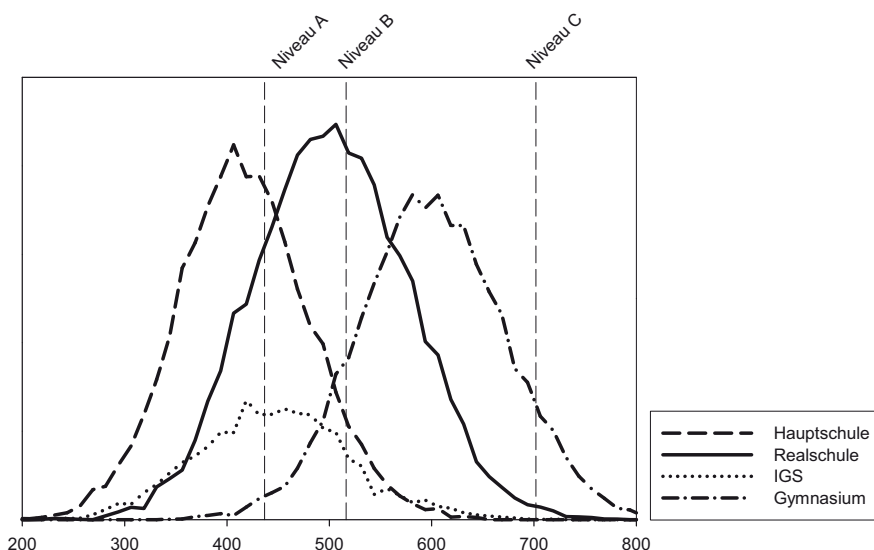


Abbildung 15.4: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprachbewusstheit Englisch Soziopragmatik zu Beginn der neunten Jahrgangsstufe.

Auf Niveau C gelingt es ihnen darüber hinaus, auch in solchen Situationen ihre Sprachbewusstheit einzusetzen, wenn die soziopragmatischen Regelungen gegebenenfalls in Konflikt zu ihrer eigenen Sprache und ihren eigenkulturellen Normen geraten oder ein tiefergehendes Erschließen erfordern.

Kompetenzniveaus Sprachbewusstheit Soziopragmatik: Verteilung auf die Bildungsgänge

Im Vergleich der Bildungsgänge fällt auf, dass sich die Kompetenzen der Schülerinnen und Schüler schwerpunktmäßig auf die Kompetenzniveaus A und B verteilen, allerdings jeweils mit sehr unterschiedlichen Anteilen (vgl. Abbildung 15.5). Dabei ist zu berücksichtigen, dass ein unterschiedlich großer Anteil von Schülerinnen und Schülern die Erfordernisse von Kompetenzniveau A noch nicht durchgängig erfüllt. In der Integrierten Gesamtschule und in der Hauptschule liegt dieser Anteil zwischen etwa 45 bzw. etwas über 60%, während er in der Realschule bei ca. 20% liegt. Insgesamt sind in der Realschule die Kompetenzniveaus A und B vorherrschend, während für das Gymnasium das Kompetenzniveau B weitgehend charakteristisch ist, wobei ein nicht zu übersehender Teil der Schülerinnen und Schüler auch das Niveau C erreichen (ca. 8%). Im Durchschnitt liegen die Leistungen in der Hauptschule mit einem Mittelwert von 416 knapp unter dem Kompetenzniveau A, an der Integrierten Gesamtschule auf Niveau A ($M = 446$), an der Realschule knapp unter Niveau B ($M = 502$), am Gymnasium in der Mitte von Kompetenzniveau B ($M = 601$).



Sprachbewusstheit Englisch Soziopragmatik zu Beginn der 9. Jahrgangsstufe

Abbildung 15.5: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprachbewusstheit Englisch Soziopragmatik zu Beginn der neunten Jahrgangsstufe nach Bildungsgängen.

Diskussion

Es ist bemerkenswert, dass Fragen der Angemessenheit von sprachlichem Verhalten von einer beachtlichen Anzahl von Schülerinnen und Schülern zutreffend bewertet werden. Gleichzeitig ist auch nicht zu übersehen, dass der Erwerb von entsprechenden Kompetenzen der Sprachbewusstheit für viele in der Fremdsprache mit großen Schwierigkeiten verbunden ist. So ist selbst ein positiver Transfer aus soziopragmatischen Vorerfahrungen – beispielsweise aus muttersprachlichen Regelungen – in vielen Fällen nicht ohne zusätzliche Lernanstrengungen zu erwarten. Es ist zu vermuten, dass viele dieser Schülerinnen und Schüler unsicher reagieren in dem Versuch, eigensprachliche soziopragmatische Normen und Regelungen in einen fremdsprachlichen Kontext zu übertragen. Darüber hinaus kann angenommen werden, dass ihre fremdsprachliche Sprachbasis noch so bruchstückhaft ist, dass das Wissen um eigensprachliche soziopragmatische Regelungen nicht angemessen aktiviert werden kann (vgl. Baker 2001, mit Bezug auf Theorien von Cummins).

15.2 Testkonstrukt Sprachbewusstheit Grammatik

Während mit der Sprachbewusstheit Soziopragmatik die Frage in den Vordergrund gestellt wird, ob fremdsprachliche Äußerungen angemessen sind, richtet das Testkonstrukt zur Sprachbewusstheit Grammatik das Augenmerk auf Fragen der

sprachlichen Korrektheit einschließlich der Fähigkeit zur Fehlerkorrektur. Dabei spielen curriculare Gesichtspunkte eine zentrale Rolle, da sowohl Formen und Strukturen berücksichtigt werden, die schon sehr früh (z.B. *Tense und Aspekt, Modalverben, Grundlagen der Wortstellung*) als auch relativ kurz vor der Durchführung des Tests (z.B. *Passiv, Nebensätze wie Relativsatz, If-Sätze, Infinitivkonstruktionen*) unterrichtet wurden. Entsprechend dem Testkonstrukt wird demnach untersucht, in welchem Ausmaß Schülerinnen und Schüler imstande sind, Verstöße gegen grammatische Regelungen des Englischen zu erkennen und gegebenenfalls zu korrigieren. Darüber hinaus wird die Fähigkeit erfasst, in einem sehr eingegrenzten Bereich fehlende grammatische Formen zu erkennen und mit Hilfe von vorgegeben Alternativen zu ergänzen. Das Sprachbewusstheitskonstrukt Englisch ist damit vorwiegend auf die Fähigkeit zur sprachlichen Fehlerkorrektur ausgerichtet. Es ist dabei weniger bedeutsam, die sprachlichen Regelungen bewusst zu durchschauen und in Regeln fassen zu können.

Kompetenzniveaus und Aufgabenbeispiele

Mit den Testaufgaben zur Sprachbewusstheit Grammatik werden vier Kompetenzniveaus zur Sprachbewusstheit Grammatik erfasst (vgl. Tabelle 15.2). Auf dem ersten Niveau (Niveau A) zeigen Schülerinnen und Schüler, dass sie einfache morphologisch-syntaktische Regelungen (z.B. *Determiners, Relativpronomen*) erkennen, vorausgesetzt es wird durch vorgegebene Auswahlmöglichkeiten deutlich auf die entsprechenden Regelungen verwiesen. Ferner können sie partiell Fehler korrigieren, wenn sich diese auf besonders vertraute Strukturen (z.B. die Unterscheidung von *who/which*) beziehen. Mit dem zweiten Niveau (Niveau B) wird ausgedrückt, dass die Schülerinnen und Schüler Verstöße gegen komplexere morphologisch-syntaktische Regelungen (z.B. Verbformen, Modalverb, Verb in indirekter Rede, Wortstellung) selbständig zu erkennen vermögen, dabei allerdings nur eine sehr eingeschränkte Fähigkeit zur Korrektur ausgeprägt haben. Die Schülerinnen und Schüler befinden sich auf Niveau C, wenn sie darüber hinaus selbständig Fehler korrigieren können. Allerdings gelingt ihnen die Korrektur auf diesem Niveau nur, wenn die entsprechenden Regelungen im Englischen und Deutschen ähnlich sind (beispielsweise Vergleichssätze oder Verben mit bestimmten syntaktischen Verknüpfungen). Erst auf Niveau D können sie auch Verstöße in Fällen korrigieren, bei denen die grammatischen Regelungen sich in den Sprachen deutlich von einander unterscheiden (z.B. bestimmte Infinitiv-Strukturen sowie Konditionalsätze, bestimmte Bereiche von *Tense und Aspect*).

Tabelle 15.2: Beschreibung der Kompetenzniveaus in Sprachbewusstheit Englisch Grammatik.

Kompetenz-niveau	Beschreibung
D (ab 805 Punkte)	Kann Fehler in komplexeren syntaktischen Regelungen sowie im Bereich von <i>Tense</i> und <i>Aspect</i> korrigieren, auch wenn die entsprechenden Regelungen der Erstsprache (gegebenenfalls der Zweitsprache) nicht parallel liegen.
C (694-804 Punkte)	Kann im Bereich komplexerer morphologisch-syntaktischer sowie syntaktischer Regelungen selbständig Fehler korrigieren, wenn die entsprechenden Regelungen der Erstsprache (gegebenenfalls auch der Zweitsprache) parallel liegen. Wenn diese Regelungen nicht parallel liegen (z.B. Comparison, Komplementierung des Verbs, <i>Tense</i> und <i>Aspect</i>), können die Fehler zwar erkannt, jedoch nur sehr partiell korrigiert werden.
B (627-693 Punkte)	Kann im Bereich komplexerer morphologisch-syntaktischer (Verbformen, Modalverb, Verb in indirekter Rede) sowie syntaktischer Regelungen (Wortstellung) selbständig Fehler erkennen, vermag sie allerdings nur sehr partiell zu korrigieren.
A (400-626 Punkte)	Kann im Bereich einfacher morphologisch-syntaktischer Regelungen (<i>Determiners</i> , Relativpronomen) Fehler erkennen, wenn deutliche Hinweise gegeben werden. Kann partiell Fehler korrigieren, wenn sich diese auf besonders vertraute Strukturen (<i>who/ which</i>) beziehen.

In den folgenden Abbildungen 15.6 bis 15.8 werden Aufgabenbeispiele für die vier Kompetenzniveaus vorgestellt. Eine detaillierte Darstellung findet sich in Nold/Rossa (2007).

Fehlt hier nicht etwas?

Vervollständige die folgenden Sätze, indem du eine der vier vorgegebenen Möglichkeiten ankreuzt. Nur eine Möglichkeit ist jeweils richtig.

I have _____ money.

- ☐ every
- ☐ a lot of
- ☐ a few
- ☐ any

Abbildung 15.6: Aufgabenbeispiel für Kompetenzniveau A in Sprachbewusstheit Englisch Grammatik.

Das Aufgabenbeispiel in Abbildung 15.6 mit seiner Lösung *a lot of* verdeutlicht, dass hier nur sehr rudimentär Sprachbewusstheit gefordert ist.

Kannst Du in den folgenden Äußerungen sprachliche Fehler entdecken?

So gehst Du vor:

Zuerst überlege, ob du in dem Kurztext der Aufgabe einen sprachlichen Fehler entdeckst. Der Fehler kann nur in einem der mit A, B oder C markierten Teile stecken.

Wenn du meinst, dass alles richtig ist, kreuze das Richtig-Kästchen an und gehe weiter zur nächsten Aufgabe.

Wenn Du einen Fehler entdeckst, kreuze eines der Kästchen A, B oder C an, um zu zeigen, wo der Fehler steckt. Dann verbessere den Fehler im Korrekturkasten.

Aufgabe:

Sue: "Justin Timberlake is coming to town. Who ☐ A want to see him?"

Harry: "We all ☐ B do. Please ☐ C don't forget to find out when he is playing!"

Abbildung 15.7: Aufgabenbeispiele für Kompetenzniveaus B und C in Sprachbewusstheit Englisch Grammatik.

Im Aufgabenbeispiel in Abbildung 15.7 ist der Fehler in A zu erkennen. Eine erfolgreiche Korrektur ist auf diesem Niveau noch nicht konsistent zu erwarten (vgl. Nold/Rossa 2007). Die Aufgabe von Kompetenzniveau B oben wird auf dem Kompetenzniveau C zusätzlich richtig korrigiert.

Sue: " ☐ A Did you hear that ? She ☐ B wants that we work ten hours today.

☐ C Isn't that crazy ?"

Abbildung 15.8: Aufgabenbeispiel für Kompetenzniveau D in Sprachbewusstheit Englisch Grammatik.

Im Aufgabenbeispiel in Abbildung 15.8 ist der Fehler in B in der gleichen Weise zu erkennen und zu korrigieren wie auf dem Kompetenzniveau C. Zusätzlich kommt hier zum Tragen, dass die Struktur *want s.o. to do ...* besonders interferenzanfällig ist.

Verteilung auf die Kompetenzniveaus Sprachbewusstheit Grammatik: Schülerleistungen insgesamt

In den folgenden Darstellungen wird deutlich, wie sich die Schülerinnen und Schüler auf die verschiedenen Niveaus im Bereich Sprachbewusstheit Englisch Grammatik verteilen.

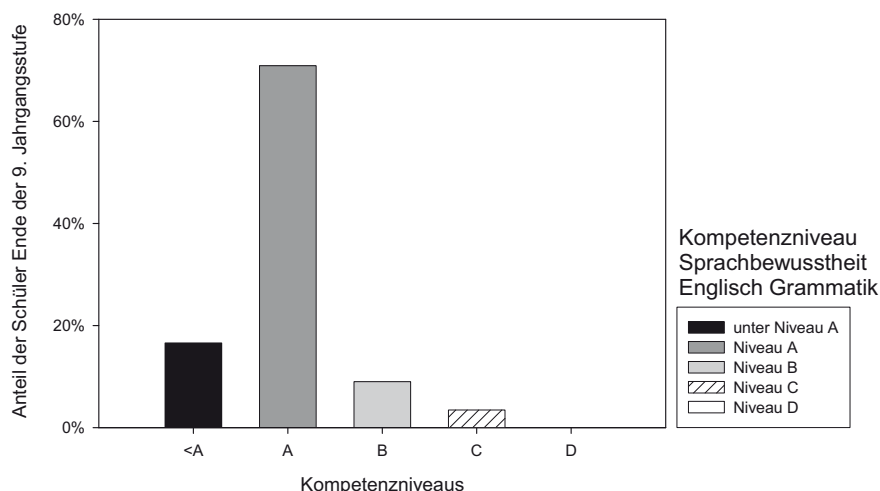


Abbildung 15.9: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprachbewusstheit Englisch Grammatik am Ende der neunten Jahrgangsstufe.

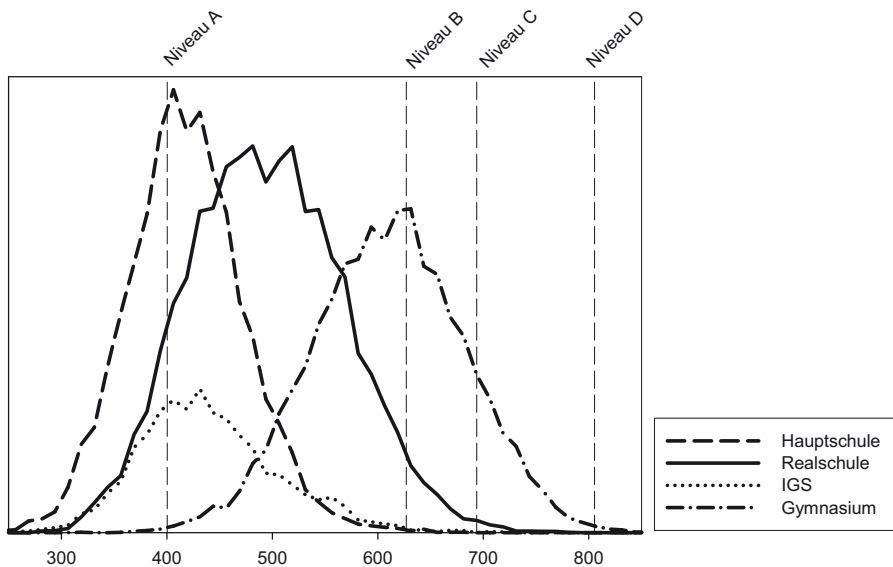
Abbildung 15.9 zeigt, dass etwas mehr als 70% der Schülerleistungen am Ende der neunten Jahrgangsstufe auf Kompetenzniveau A liegen; ein Anteil von ungefähr 16% der Schülerleistungen liegt unter diesem Niveau. Die Schülerinnen und Schüler auf Niveau A sind in der Lage, mit sehr elementaren grammatischen Regelungen umzugehen, vorausgesetzt es werden Lösungsmöglichkeiten vorgegeben oder die einfachen grammatischen Regelungen sind intensiv im Unterricht behandelt worden. Sie besitzen damit ein eingeschränktes Sprachgefühl in der Grammatik des Englischen. Auf Niveau B liegen etwa neun Prozent der Schülerleistungen, weitere drei bis vier Prozent auf Niveau C. Schülerinnen und Schüler dieser Niveaus verfügen über eine deutlich begrenzte bis etwas stärker entwickelte Sprachbewusstheit in der englischen Grammatik. Sie sind fähig, in bestimmten grammatischen Bereichen sprachliche Verstöße zu entdecken und sie je nach Niveau in unterschiedlichem Maße zu korrigieren. Nur eine verschwindend kleine Gruppe erreicht eine darüber hinaus gehende Sprachbewusstheit (Niveau D) in der englischen Grammatik.

Verteilung auf die Kompetenzniveaus Sprachbewusstheit Grammatik: Unterschiede zwischen den Bildungsgängen

Schon bei der Analyse der Schülerleistungen insgesamt ist nicht zu übersehen, dass das erste Kompetenzniveau der Sprachbewusstheit Grammatik eine herausragende Rolle spielt (vgl. Abbildung 15.10). Es überrascht daher nicht, dass dieses Niveau auch für die verschiedenen Bildungsgänge prägend ist. So befinden sich selbst im Gymnasium fast 60% der Schülerinnen und Schüler auf diesem Niveau. In der Realschule sind dies ca. 87%, in der Integrierten Gesamtschule und der Hauptschule ungefähr 70% bzw. 62%. Der im Vergleich zur Realschule niedrigere Anteil in der

Integrierten Gesamtschule und Hauptschule erklärt sich dadurch, dass in den beiden Bildungsgängen der Anteil unter Niveau A vergleichsweise größer ist (fast 29% bzw. 38%). Ausschließlich im Gymnasium und der Realschule finden sich Schülerinnen und Schüler, die ein höheres Niveau erreichen. In der Realschule liegt dieser Prozentsatz etwas unter vier Prozent, wobei hier Niveau B deutlich überwiegt. Im Gymnasium erreichen fast 41% der Schülerinnen und Schüler höhere Niveaus. Im Wesentlichen teilen sie sich in zwei Gruppen auf: Etwas über 28% sind Niveau B zuzurechnen, 12% befinden sich auf Niveau C. Nur 0.2% erreichen Niveau D der Sprachbewusstheit Grammatik.

Im Durchschnitt liegen die Leistungen in allen Bildungsgängen auf Kompetenzniveau A. An der Hauptschule ist ein Mittelwert von 419 zu verzeichnen, an der Gesamtschule ein Mittelwert von 441, an der Realschule ein Wert von 494, am Gymnasium werden im Mittel 610 Punkte auf der Skala erreicht (vgl. Abbildung 15.10).



Sprachbewusstheit Englisch Grammatik Ende der 9. Jahrgangsstufe

Abbildung 15.10: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprachbewusstheit Englisch Grammatik am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Diskussion

Die Ergebnisse in den Bereichen Sprachbewusstheit Soziopragmatik und Grammatik unterscheiden sich unübersehbar. Dies ist zunächst nicht verwunderlich, da in der Planung der DESI-Tests Sprachbewusstheit Englisch von zwei deutlich unterscheidbaren Testkonstrukten ausgegangen wurde. Dennoch sind die Unterschiede im Einzelnen bemerkenswert. So teilen sich die Schülerinnen und Schüler im Bereich

der Sprachbewusstheit Soziopragmatik grob gesehen in drei Gruppen auf. Zwei von ihnen haben Kompetenzen entwickelt, die den Niveaus A und B entsprechen, während die dritte Gruppe diese Niveaus nicht erreicht. Werden die Bildungsgänge in diese Überlegungen einbezogen, so zeigt sich, dass diese drei Gruppen sich verstärkt in gegenläufiger Richtung auf die Bildungsgänge verteilen. Besonders auffallend ist dabei der starke Kontrast zwischen den Kompetenzen in der Hauptschule und dem Gymnasium. Dagegen befindet sich im Bereich der Sprachbewusstheit Grammatik die überwiegende Mehrzahl der Schülerinnen und Schüler auf Niveau A, wobei sich die Anteile auf diesem Niveau in der Hauptschule und dem Gymnasium nur unwesentlich unterscheiden. Darüber hinaus fällt allerdings auf, dass auch hier eine gegenläufige Richtung insofern zu beobachten ist, als in der Hauptschule ein beachtlicher Anteil unter Niveau A liegt, während im Gymnasium ein entsprechender Anteil erwartungsgemäß über diesem Niveau liegt. Die anderen Bildungsgänge unterscheiden sich hier weniger deutlich.

Eine mögliche Ursache für die in allen Bildungsgängen gleichmäßig gering ausgebildete Sprachbewusstheit Grammatik ist unter anderem vermutlich darin zu sehen, dass sprachliche Korrekturen im Englischunterricht generell vorwiegend als eine Aufgabe der Lehrenden verstanden werden und somit Selbstkorrektur oder die Verbesserung durch die Lernenden eher die Ausnahme darstellen. Untersuchungen zum sprachlichen Korrekturverhalten (vgl. Andrews 1997; Leeman 2003; Lyster/Ranta 1997; Tedick/de Gortari 1998) haben jedoch gezeigt, dass insbesondere die Selbstkorrektur, an zweiter Stelle auch die Korrektur durch Mitschülerinnen und Mitschüler sich positiv auf die Sprachentwicklung auswirkt. Einer der Hauptgründe dafür wird darin gesehen, dass die Lehrerkorrekturen nicht als spezifische Verbesserungsvorschläge wahrgenommen werden. Auch bleibt zu fragen, ob die Schülerinnen und Schüler genügend Möglichkeiten haben, um eigengesteuert sprachliche Formulierung zu leisten.

Für den Bereich der Soziopragmatik ist zu untersuchen, ob im Englischunterricht ausreichend Gelegenheit gegeben wird, um kulturspezifische Verhaltensweisen, die sich im Diskurs manifestieren, bewusst wahrzunehmen und sie in Auseinandersetzung mit eigenkulturellen Normen stärker zu begreifen (vgl. Kasper/Rose 2002).

Literatur

- Alderson, J./Clapham, C./Steel, D. (1997): Metalinguistic knowledge proficiency, language aptitude and language proficiency. In: *Language Teaching Research* 1/2, S. 93-121.
- Andrews, S. (1997): Metalinguistic awareness and teacher explanation. In: *Language Awareness* 6, H. 2/3, S. 147-61.
- Aplin, R. (1997): *Knowledge about language and language awareness: an annotated bibliography*. Leicester: Leicester University School of Education.
- Baker, C. (2001): *Foundations of Bilingual Education and Bilingualism*. Clevedon, Buffalo, Toronto, Sydney: Multilingual Matters.
- Beck, B. /Klieme E. (Hrsg.) (2007): **Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie** (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Cameron, L. (1993): Degrees of knowing. In: *Language Awareness* 2, H. 1, S. 3-13.

- Düwell, H./Gnutzmann, C./Königs, F.G. (Hrsg.) (2000): Dimensionen der Didaktischen Grammatik. Festschrift für Günther Zimmermann zum 65. Geburtstag. Bochum: AKS-Verlag.
- Eichler, W./Nold, G. (2007): Sprachbewusstheit. In: Beck, B. /Klieme E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 59-78.
- Ellis, N. (Hrsg.) (1994a): Implicit and explicit learning of languages. London: Academic Press.
- Ellis, R. (1994b): The study of Second Language Acquisition. Oxford: Oxford University Press.
- Ellis, R. (2004): The definition and measurement of L2 explicit knowledge. In: Language Learning 54, H. 2, S. 227-275.
- Han, Y./Ellis, R. (1998): Implicit knowledge, explicit knowledge and general language proficiency. In: Language Teaching Research 2, H. 1, S. 1-23.
- Hawkins, E.W. (1999): Foreign Language Study and Language Awareness. In: Language Awareness 3&4, S. 124-142.
- Karmiloff-Smith, L. (1992): Beyond modularity. A developmental perspective on cognitive science. Cambridge, MA: MIT Press.
- Kasper, G./Rose, K.R. (2002): Pragmatic development in a second language. Oxford: Blackwell.
- Leeman, J. (2003): Recasts and second language development. In: Studies in Second Language Acquisition 25, S. 37-63.
- Lyster, R./Ranta, L. (1997): Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. In: Studies in Second Language Acquisition 19, S. 37-66.
- Nold, G./Grimmig, R. (2000): Zum Sprachgefühl und Sprachwissen von Fremdsprachenlernern im Bereich der Grammatik des Englischen. In: Düwell, H./Gnutzmann, C./Königs, F.G. (Hrsg.): Dimensionen der Didaktischen Grammatik. Festschrift für Günther Zimmermann zum 65. Geburtstag. Bochum: AKS-Verlag, S. 211-228.
- Nold, G./Rossa, H. (2007). Sprachbewusstheit. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz, S. 222-240.
- Paradis, M. (2004): Neurolinguistics of bilingualism and the teaching of languages. URL: http://semtoticon.com/virtuals/talks/paradis_txt.htm (Zugriff am 27.05.2004).
- Schmidt, R. (Hrsg.) (1995): Attention and Awareness in Foreign Language Learning. Honolulu: University of Hawaii Press.
- Tedick, D.J./de Gortari, B. (1998): Research on Error correction and implications for classroom teaching. In: Acie Newsletter vol. 1, No. 3, S. 1-6. URL: <http://www.carla.umn.edu/immersion/acie/vol1/May1998.pdf> (Zugriff 05. 12. 2006).

Günter Nold / Henning Rossa

16 Sprechen Englisch

Die mündliche Sprachproduktion von Fremdsprachenlernern kann in Testsituationen entweder in Nachbildung von realen Gesprächssituationen kommunikativ modelliert oder im Sinne von psycholinguistisch definierten Teilkompetenzen der mündlichen Sprachkompetenz erfasst werden. Im letzteren Fall werden auf der Basis von Indikatoren dieser Teilkompetenzen die sprachproduktiven Fähigkeiten eingeschätzt (vgl. Luoma 2004).

Einerseits wird somit in Annäherung an eine reale Gesprächssituation ein gesprochener Text entwickelt, der als ein Ausdruck der sich entwickelnden mündlichen Sprachkompetenz von Schülerinnen und Schülern betrachtet wird. Das entsprechende Sprachprodukt ermöglicht es zu analysieren, welche Redemittel für bestimmte Sprechakte im Diskurs thematisch, sprachlich und situationsbezogen verfügbar sind. Bei Abwägung der Vor- und Nachteile einer solchen Modellierung von Kommunikation ist jedoch nicht zu übersehen, dass die Nachbildung einer realen Gesprächssituation zusätzlich zu der Testperson den Einsatz von zumindest einem weiteren Gesprächspartner (einer Gesprächspartnerin) erfordert (vgl. Brown 2003). Dadurch wird es unumgänglich, die Gesprächsleistung der Testperson auch in Relation zu den Besonderheiten des Gesprächspartners zu bewerten.

Andererseits können sprachliche Komponenten der mündlichen Sprachkompetenz in den Mittelpunkt eines Tests zum Sprechen gestellt werden. Luoma prägt für die entsprechende testtheoretische Vorgehensweise den Begriff *structured speaking task* (Luoma 2004, S. 50, 158 ff). Die sprachlichen Komponenten werden aus einer psycholinguistischen Perspektive als Teilfähigkeiten verstanden, die für das Sprechen grundlegend sind. Die zu den sprachlichen Komponenten gewonnenen Daten geben als indirekte Indikatoren Aufschluss darüber, wie ausgeprägt die mündlich-produktive Kompetenz einer Testperson ist, um angemessen und effektiv sprechen zu können.

Im DESI-Test zum Sprechen wird sowohl ein kommunikativer als auch ein psycholinguistisch orientierter Weg beschritten. So werden einerseits kommunikative Gesprächselemente mit sowohl eng geführten als auch offenen Antwortmöglichkeiten verwendet und andererseits die sprachlichen Teilkomponenten der mündlich-produktiven Kommunikationskompetenz zum Gegenstand von Testaufgaben gemacht. Damit liegen den Testaufgaben einerseits ein kommunikatives Testkonzept und andererseits ein psycholinguistisches Modell der Sprechfähigkeit zu Grunde (vgl. Levelt 1989; für Teilkompetenzen auch Buck 2001; Ellis 2002; Kramsch 1986; Read 2000; Vollmer u.a. 2001).

16.1 Konstrukt SET-10 Test

Die Kompetenzen im Bereich des Sprechens werden mit dem computergesteuerten Testprogramm des SET-10 Tests (Ordinate, www.ordinate.com) erfasst. Der mehr kommunikativ orientierte Testteil erfasst vor allem die Fähigkeit, einerseits auf inhaltlich eng geführte Kurzfragen zutreffende Antworten geben zu können und andererseits auf offene Fragen mit einer inhaltlich angemessenen freien Antwort reagieren zu können. Die stärker psycholinguistisch orientierten Testaufgaben konfrontieren die Testpersonen mit konkreten Sprechanforderungen, die von ihnen in Echtzeit hörend und sprechend, jedoch ohne einen Kontextbezug zu bewältigen sind. Mit ihnen lässt sich prognostizieren, in welchem Ausmaß die Testpersonen fähig sind, sich mit Sprechern des Englischen in normaler Sprechgeschwindigkeit verständlich zu machen und inhaltlich zusammenhängend zu verständigen (vgl. Nold/DeJong 2007).

Im Testkonstrukt Sprechen werden dementsprechend folgende Fähigkeiten des Sprechens direkt und indirekt erfasst:

“The SET-10 test measures facility in spoken English – that is, the ability to understand spoken English on everyday topics and to respond appropriately at a native-like conversational pace in intelligible English. Another way to express the construct facility in spoken English is ‘ease and immediacy in understanding and producing appropriate conversational English.’ This definition relates to what occurs during the course of a spoken conversation. While keeping up with the conversational pace, a person has to track what is being said, extract meaning as speech continues, and then, on occasion, formulate and produce a relevant and intelligible response.” (Ordinate 2004, S. 2)

Im Einzelnen bestimmen folgende Aspekte das Testkonstrukt:

- Fragen von Sprechern/Sprecherinnen über Alltagsthemen verstehen und mit Meinungsäußerungen angemessen beantworten können;
- auf den Sprechimpuls eines Sprechers/einer Sprecherin im Gespräch in kurzer Zeit inhaltlich zutreffend und kohärent reagieren können, und zwar mit normaler Sprechgeschwindigkeit;
- das eigene Verständnis einer gehörten Äußerung durch eine flüssige Wiederholung der Äußerung mitteilen oder durch eine zutreffende Handlung signalisieren können;
- einen Text verständlich und phonetisch-phonologisch korrekt und flüssig vorlesen können;
- Lexik und Grammatik als sprachliche Redemittel effektiv einsetzen können;
- Äußerungen phonetisch-phonologisch korrekt und flüssig formulieren können.

Einerseits stehen somit die Mitteilungsfunktion von Sprache und der Bezug auf alltagssprachliche Themen im Mittelpunkt, andererseits werden sprachliche Daten als indirekte Indikatoren für die mündliche Sprachkompetenz erfasst. In kommunikativer Hinsicht wird ein kohärenter interaktiver Text entworfen. In Ansätzen wird auch eine Adaptationsfähigkeit im Rahmen dieser Kompetenz verlangt (vgl. Johnstone 2002),

insofern über die Äußerungsgrenze eines Sprechers die Kohärenz gewahrt werden muss. Nur die Initiative zum Sprecherwechsel und zum Bestimmen des Themas liegt nicht in der Hand der Testteilnehmer.

Konkret wurden diese Aspekte des Testkonstrukts im SET-10 Test in der folgenden Weise operationalisiert: Der Test besteht aus fünf Teilen mit insgesamt 53 Testaufgaben, die für jede Testperson aus einem Datenpool zusammengestellt werden. Jeder Testteil konfrontiert die Probanden mit einer spezifischen Aufgabe (*Test Task*). Die sprachlichen Anstöße klingen dabei natürlich, da sie auf Aufnahmen von realen menschlichen Stimmen beruhen. Die Probanden werden einzeln für ca. zehn Minuten in ein Telefongespräch verwickelt. In diesem Gespräch befolgen sie mündliche Anweisungen. Sie lesen acht Sätze aus einer Gruppe von zwölf ihnen schriftlich vorliegenden Sätzen nach Anweisung laut vor, wiederholen 16 unterschiedlich komplexe Äußerungen wörtlich (auf dem Hintergrund einer Aktivierung von mehr oder weniger automatisierten Redeteilen), geben 16 Kurzantworten auf wirkliche Fragen zu vertrauten Alltagsbegebenheiten und bringen ferner mündlich zehn syntaktisch ungeordnete Redeteile in eine syntaktisch richtige Abfolge. Schließlich beantworten sie drei Fragen mit freien Antworten in einer festgelegten Zeitspanne von jeweils 20 Sekunden. Die Fragen dieses Testteils wurden in Abwandlung des ursprünglichen Tests sowohl in ihrer inhaltlichen als auch sprachlichen Komplexität speziell auf die Besonderheiten der Altersgruppe der DESI-Studie und auf die zu erwartenden Kompetenzniveaus des GERS zugeschnitten.

Die Testaufgaben fokussieren damit trotz aller Künstlichkeit einer Testsituation folgende Aspekte der mündlichen Sprechfähigkeit im Englischen:

- offene Fragen angemessen kommentieren,
- zutreffende Kurzantworten geben,
- Verständnis signalisieren,
- Verfügbarkeit von lexikalischen Redemitteln,
- Aktivierung von mehr oder weniger automatisierten grammatischen Redemitteln,
- phonetisch-phonologisch korrekte Artikulation und
- Flüssigkeit der Rede.

Das computergesteuerte Testprogramm errechnet für die Testteile, die eindeutig vorhersagbare Elemente enthalten, Einschätzungen der Testleistungen in den Teilkomponenten Aussprache, Flüssigkeit der Rede, Wortschatz und Satzbau (*subscores*) und erstellt zusätzlich eine Gesamteinschätzung (*overall score*). Diese Einschätzungen sind auf die Niveaus der Sprechfähigkeit des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GERS) abgestimmt (De Jong/Bernstein 2001; Europarat 2001). Die Werte des SET-10 Tests, die den Übergang von einem zum nächst höheren Niveau (*cut-off points*) markieren, erscheinen in der Beschreibung der Kompetenzniveaus unten.

Der kommunikative Testteil, der aus Fragen mit offenen Antwortmöglichkeiten (*Task*: Fragen möglichst ausführlich beantworten) besteht, ergänzt den stärker psycholinguistisch orientierten Testanteil und validiert die automatische

Leistungseinschätzung des Computerprogramms, indem auf der Basis der Niveaus des GERS ein eigenes Ratingverfahren von Experten durchgeführt wird (vgl. Nold/DeJong 2007), um die freien Antworten global einzuschätzen. Facetten der Diskurskompetenz sind damit deutlich auch direkt im Test repräsentiert.

Zusätzlich wurde das Testformat des SET-10 Tests validiert, indem in einer Interventionsstudie mit drei Klassen der neunten Jahrgangsstufe untersucht wurde, ob ein Training des Testformats einen Einfluss auf das Testergebnis hat. Die Ergebnisse ($n = 60$, zwei Testzeitpunkte mit einer Intervention zwischen den Testzeitpunkten 1 und 2) zeigen, dass nach einem Testtraining im Sinne von *teaching to the test* ein statistisch signifikanter Einfluss auf die Testergebnisse sich nicht nachweisen lässt (vgl. Pilotstudie von Hinz/Papenberg/Rossa/Nold 2006).

16.2 Kompetenzniveaus (Deskriptoren)

Die Kompetenzniveaus beim Sprechen werden entsprechend dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (Europarat 2001) als einem Außenkriterium bestimmt. Die curricularen Bestimmungen der Länder können dabei als weitgehend berücksichtigt betrachtet werden. Mithilfe des SET-10 Tests ist es möglich, den Schülerinnen und Schülern bestimmte Kompetenzniveaus hinsichtlich ihrer Sprechfähigkeit in Englisch zuzuordnen (vgl. Tabelle 16.1).

16.3 Stichprobe SET-10 Test

Anders als in der Gesamtstichprobe von DESI wurde der SET-10 Test aus Gründen der Machbarkeit nur in einer Teilstichprobe von insgesamt 1870 Schülerinnen und Schülern durchgeführt. Dabei wurden aus jeder Klasse jeweils drei Mädchen und drei Jungen nach dem Zufallsprinzip für die Testteilnahme ausgewählt. Von diesen Schülerinnen und Schülern liegen dementsprechend Testdaten vor. Alle verfügbaren Schülerdaten wurden im Rahmen der computergesteuerten Auswertung in Hinsicht auf die Kompetenzniveaus ausgewertet. Aus den *subscores* für Aussprache, Flüssigkeit, Wortschatz und Satzbau wurde dabei das *overall score* gebildet. Darüber hinaus wurden die offenen Items (jeweils drei Fragen mit offenen Antwortmöglichkeiten für jede Testperson) in einem speziellen Ratingverfahren analysiert. Bei diesem Ratingverfahren wurde der Qualität der Auswertung eine höhere Priorität eingeräumt als der Vollständigkeit der Datenanalyse, wie aus der folgenden Darstellung ersichtlich ist.

Tabelle 16.1: Beschreibung der Kompetenzniveaus in Sprechen Englisch.

Kompetenz-niveau	Beschreibung
C2 (79-80 Punkte)	Kann klar, flüssig und gut strukturiert sprechen und seinen Beitrag so logisch aufbauen, dass es den Zuhörern erleichtert wird, wichtige Punkte wahrzunehmen und zu behalten.
C1 (69-78 Punkte)	Kann komplexe Sachverhalte klar und detailliert beschreiben und darstellen und dabei untergeordnete Themen integrieren, bestimmte Punkte genauer ausführen und alles mit einem angemessenen Schluss abrunden.
B2 (58-68 Punkte)	Kann zu einer großen Bandbreite von Themen aus seinen/ihren Interessensgebieten klare und detaillierte Beschreibungen und Darstellungen geben, Ideen ausführen und durch untergeordnete Punkte und relevante Beispiele stützen.
B1 (47-57 Punkte)	Kann Sachverhalte klar und systematisch beschreiben und darstellen und dabei wichtige Punkte und relevante stützende Details angemessen hervorheben. Kann relativ flüssig eine unkomplizierte, aber zusammenhängende Beschreibung zu Themen aus seinen/ihren Interessengebieten geben, wobei die einzelnen Punkte linear aneinander gereiht werden.
A2 (36-46 Punkte)	Kann eine einfache Beschreibung von Menschen, Lebens- oder Arbeitsbedingungen, Alltagsroutinen, Vorlieben oder Abneigungen usw. geben, und zwar in kurzen, listenhaften Abfolgen aus einfachen Wendungen und Sätzen.
A1 (26 -35 Punkte)	Kann sich mit einfachen, überwiegend isolierten Wendungen über Menschen und Orte äußern.
Unter A1 (< 25 Punkte)	Ungenügende Grundlagen für eine Bewertung, z.B. Schweigen, einzelne Wörter ohne erkennbare Bedeutung, irrelevante Antworten, etc.

16.4 Rating der offenen Aufgaben und Bezug zum computergesteuerten Rating

Ausgewählte Ergebnisse der offenen Aufgaben wurden durch ein Team von sechs Ratern eingeschätzt, die ein Ratertraining auf Basis der Kompetenzniveaus der mündlichen Sprachkompetenz durchlaufen hatten. Im Rahmen einer Zufallsstichprobe wurden von je einem Schüler/einer Schülerin aus jeder Klasse der DESI-Stichprobe die SET-10 Testdaten der offenen Items bewertet. D.h. jeweils drei offene Items von insgesamt 361 Schülerinnen und Schülern wurden in der Regel von jeweils fünf Ratern geschätzt. Jedes Item wurde dabei mindestens von zwei verschiedenen Ratern bewertet. Zusätzlich wurden zehn Prozent der Items von jedem Rater nach dem Zufallsprinzip im Prozess der Bewertung doppelt eingeschätzt, um die Verlässlichkeit der Einschätzung eines jeden Raters zu prüfen (im Sinne der Feststellung der Intra-Rater-Reliabilität). Insgesamt wurden auf diese Weise für 1320 offene Items (einschließlich der ca. 10% doppelten Items) 2605 Einschätzungen bezogen auf die

Kompetenzniveaus festgelegt. Dabei war die Inter-Rater-Reliabilität mit $r = .89$ angemessen.

Die Korrelation zwischen diesen Rating-Werten der offenen Items und dem vom Computer ermittelten Gesamtergebnis der geschlossenen Items (*overall score*) ist $r = .76$. Die Teilkompetenzen - erfasst in den *subscores* des SET-10Tests - korrelieren unterschiedlich hoch mit den Rating-Werten:

- Beherrschung des Satzbaus $r = .71$,
- Wortschatz $r = .59$,
- Flüssigkeit $r = .61$,
- Aussprache $r = .65$.

Bei der Bewertung dieser Korrelationen ist zu berücksichtigen, dass den zwei Arten von Werten unterschiedliche Teile des Testkonstrukts zu Grunde liegen. Es ist daher auch nicht unerwartet, dass die Komponenten (*subscores*) des SET 10-Tests unterschiedlich hoch mit den Werten der offenen Items korrelieren. Es bestätigt sich damit, dass mit dem Test deutlich unterscheidbare Facetten der Kompetenz des Sprechens erfasst werden. Insgesamt ist festzustellen, dass die SET-10 Testwerte verlässliche und valide Daten darüber zur Verfügung stellen, welche Kompetenzniveaus die DESI-Schülerinnen und Schüler im Bereich der mündlichen Sprachproduktion erreicht haben.

16.5 SET-10 Testergebnisse

Verteilungen insgesamt

Die Testleistungen der Schülerinnen und Schüler – gestützt auf die Werte der automatischen Auswertung der Schülerdaten – verteilen sich am Ende der neunten Jahrgangstufe in folgender Weise auf die Kompetenzniveaus (vgl. Abbildung 16.1): Etwas mehr als ein Drittel aller Schülerinnen und Schüler erreicht schon am Ende der neunten Jahrgangstufe mindestens das Niveau B1 des GERS, das curricular für das Ende der Sekundarstufe I vorgesehen ist. Unter ihnen gibt es eine Gruppe von etwa zehn Prozent Schülerinnen und Schülern, die bereits am Ende der Jahrgangstufe Neun das für das Abitur vorgesehene Niveau B2 in der Sprechfähigkeit erreichen oder sogar darüber hinausgehen. Ein knappes weiteres Drittel der Schülerinnen und Schüler befinden sich auf dem Niveau A2 des GERS; sie erreichen damit zumindest das für den Abschluss des Hauptschulstudiengangs vorgesehene Niveau. Die restlichen Schülerinnen und Schüler, etwa ein Drittel, sind am Ende der neunten Jahrgangstufe noch nicht imstande, in der Sprechfähigkeit über sehr basale und eingeschränkte Grundfähigkeiten hinauszugelangen. Innerhalb dieses Drittels gibt es darüber hinaus eine kleinere Gruppe, der es im Englischunterricht nicht gelingt, die Fremdsprache über einzelne angelernte Wendungen hinaus kreativ zu verwenden (unter Niveau A1).

Insgesamt ist im Stil einer Gaußkurve ersichtlich, dass die verschiedenen Niveaus des GERS bis einschließlich C vertreten sind. Beachtenswert ist dabei die große

Heterogenität der Sprechfähigkeit, die sich hier vor allem im Kontrast zwischen den Niveaus A1 und darunter im Gegensatz zu B und C zeigt und die damit in hohem Maße den Verteilungen in den anderen Tests im Englischteil von DESI entspricht, auch wenn beispielsweise in Hör- und Leseverstehen (vgl. Kapitel 11 und 12) im unteren Kompetenzbereich nicht so deutlich differenziert werden konnte wie im Falle des SET-10 Tests.

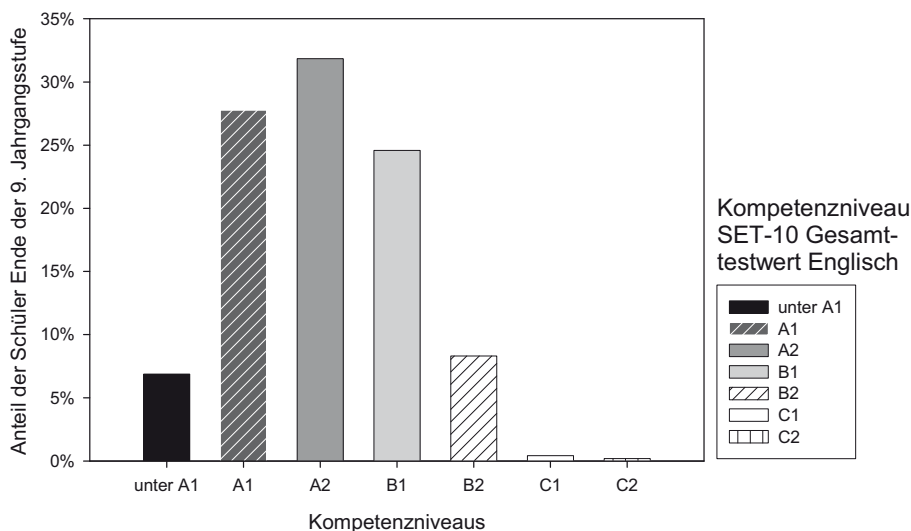


Abbildung 16.1: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprechen Englisch am Ende der neunten Jahrgangsstufe.

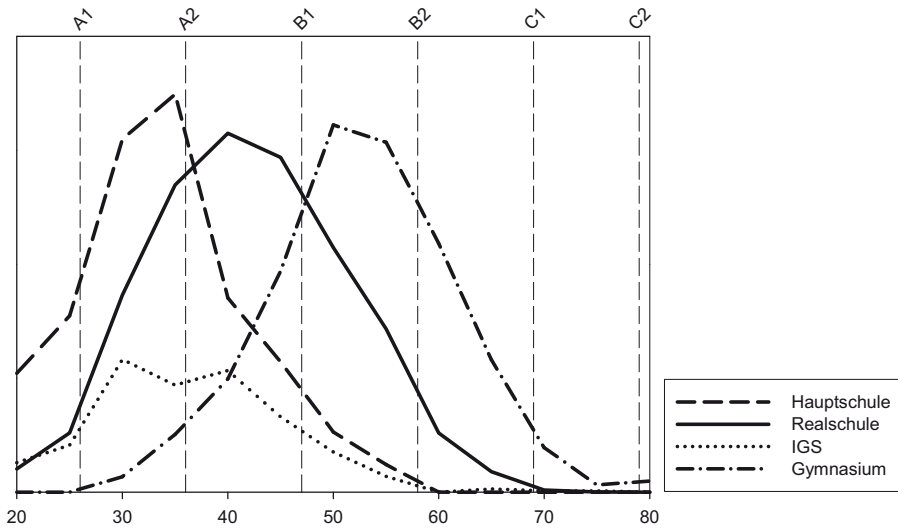
Verteilungen bezogen auf die Bildungsgänge

Die Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in den verschiedenen Bildungsgängen ist in der nachfolgenden Abbildung 16.2 dargestellt. In der Verteilung auf die Bildungsgänge lassen sich in den übrigen Englischtests vergleichbare Tendenzen erkennen. Beachtenswert ist der Kontrast zwischen dem Hauptschul-Bildungsgang und der Integrierten Gesamtschule einerseits, sowie dem gymnasialen Bildungsgang andererseits, wobei die Schülerinnen und Schüler des Realschul-Bildungsgangs in der Mitte liegen.

Innerhalb des Hauptschul-Bildungsgangs selbst erreicht etwas mehr als ein Drittel der Schülerinnen und Schüler das Niveau A2 und teilweise B1 des GERS. Der SET-10 Test erfasst jedoch im Gegensatz zu den DESI-Tests zum Hör- und Leseverstehen auch Fähigkeiten unter diesem curricularen Niveau. Damit wird deutlich, dass jeder zweite Hauptschüler dem Niveau A1 zugeordnet werden kann; die Schülerinnen und Schüler können sich also zumindest in einfachen Wendungen mündlich äußern; 16% der Hauptschülerinnen und Hauptschüler bleiben unter diesem Niveau.

Das Niveau B1, das laut KMK-Bildungsstandards den mittleren Schulabschluss am Ende der Jahrgangsstufe Zehn charakterisiert, wird ein Jahr vor diesem Bezugs-

datum bereits von fast 28% der Schülerinnen und Schüler im Bildungsgang Realschule und von 74% im Gymnasialbereich erreicht oder übertroffen. Besondere Erwähnung sollte der Anteil von Schülerinnen und Schülern in sehr hohen Niveaus bereits in der Jahrgangsstufe Neun finden.



SET-10 Gesamtttestwert Ende der 9. Jahrgangsstufe

Abbildung 16.2: Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprechen Englisch am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

Die genauen Prozentsätze der Leistungsverteilung *Sprechen Englisch* sind in Tabelle 16.2 zusammengestellt.

Tabelle 16.2: Prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzniveaus in Sprechen Englisch am Ende der neunten Jahrgangsstufe nach Bildungsgängen.

	Gesamt	Hauptschule	Realschule	IGS	Gymnasium
Niveau	%	%	%	%	%
unter A1	6.9	16.3	3.5	12.1	-
A1	27.8	49.1	26.8	41.0	4.1
A2	31.8	28.6	42.0	36.2	21.9
B1	24.6	6.0	24.0	9.7	47.9
B2	8.3	-	3.6	0.5	24.1
C1	0.4	-	-	0.4	1.3
C2	0.2	-	-	-	0.7

Verteilungen nach Geschlecht

Wenn die Geschlechterverteilung im Rahmen des SET-10 Tests analysiert wird, fällt auf, dass die Jungen in diesem Test in ihren Kompetenzen etwas über denen der Mädchen liegen (vgl. Abbildung 16.3). Ob hier das Testformat eine für die Jungen positive Rolle spielt oder ob sich in den Werten besondere Unterrichtsmerkmale wie der Sprechanteil der Schüler im Vergleich zu dem der Schülerinnen widerspiegelt, lässt sich ohne weitere Analysen nicht klären. Hier sind Untersuchungen anzustellen, die über den Untersuchungsgegenstand der Studie zum Testformat von Hinz u.a. (2006) hinausgehen. Mit Hilfe der in DESI erfassten Unterrichtsdaten sowie Hintergrundvariablen wird es möglich sein, diesen Befund differenzierter zu betrachten.

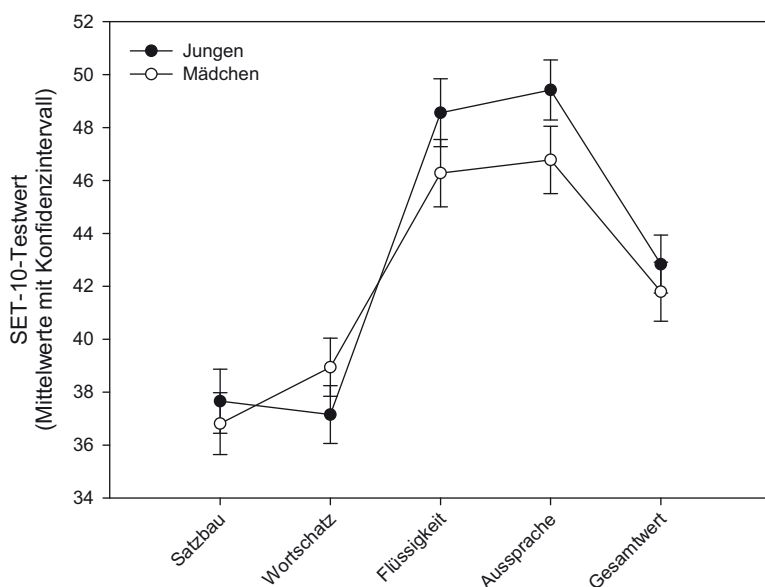


Abbildung 16.3: Mittlere Testleistungen in den einzelnen SET-10 Test-Beurteilungskriterien nach Geschlecht.

16.6 Diskussion

Die Ergebnisse des SET-10 Tests sind in ihrer Verteilung auf die Kompetenzniveaus und die Bildungsgänge weitgehend den Ergebnissen in den Hör-, Leseverstehenstests und dem C-Test vergleichbar, auch wenn erwartungsgemäß Unterschiede in den Einzelheiten unverkennbar sind (vgl. Oller 1983 und die Diskussionen über die Struktur der Sprachkompetenz oder -kompetenzen). Da der SET-10 Test empirisch nachgewiesenermaßen an die Niveaus des Gemeinsamen Europäischen Referenzrahmens für Sprachen (Europarat 2001) gebunden ist, stellt er in der DESI-Studie eine unmittelbare Brücke zum GERS dar. Die SET-10 Testergebnisse beruhen zwar auf einem

spezifischen und zugleich originellen Testkonstrukt zur Sprechfähigkeit; indirekt lassen sie dennoch darauf schließen, dass beispielsweise unterhalb des Niveaus A2 insbesondere in den Hör- und Leseverstehenstests ein breites A1-Niveau angenommen werden kann, das hier im SET-10 Test deutlich differenziert ausgewiesen erscheint. Allerdings entspricht dieses A1 Niveau wie auch in den anderen Tests nicht den angestrebten Anforderungen der Curricula.

Es ist ferner bemerkenswert, dass das computerbasierte Testformat des SET-10 Tests sehr positive Reaktionen von Seiten der Schülerinnen und Schüler hervorrief. Diese Erfahrung ermutigt dazu, in Sprachtests neue Wege auch unter Einsatz von computergesteuerten Systemen zu wagen.

Literatur

- Beck, B./Klieme, E. (Hrsg.) (2007): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Brown, A. (2003): Interviewer variation and the co-construction of speaking proficiency. In: Language Testing 20, S. 1-25.
- Buck, G. (2001): Assessing Listening. Cambridge: CUP.
- Ellis, N. (2002): Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. In: Studies in Second Language Acquisition 24, S. 143-188.
- De Jong, J. H. A. L./Bernstein, J. (2001): Relating PhonePass^(tm) Reported Scores to the Council of Europe Framework Level Descriptors. Proceedings of Eurospeech 2001. URL <http://lc.ust.hk/~centre/conf2001/proceed/dejong.pdf> (Zugriff im März 2007).
- Europarat (Hrsg.) (2001): Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen. Berlin: Langenscheidt.
- Hinz, S./Papenberg, S./Rossa, H./Nold, G. (2006): Pilotstudie zu SET 10-Test. Vortrag auf der DESI-Arbeitstagung an der Johann Wolfgang Goethe-Universität Frankfurt/Main am 12.-13.5.2006.
- Johnstone, B. (2002): Discourse Analysis. Malden, Oxford: Blackwell Publishers.
- Kramsch, C. (1986): From language proficiency to interactional competence. In: The Modern Language Journal 70, H. 4, S. 366-372.
- Luoma, S. (2004): Assessing Speaking. Cambridge: Cambridge University Press.
- Levelt, W. J. M. (1989): Speaking: from intention to articulation. Cambridge, Mass.: MIT Press.
- Nold, G./De Jong, J. H. A. L. (2007): Sprechen. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz., S. 241-251.
- Oller, J. W. Jr. (1983): Evidence for a general language proficiency factor: an expectancy grammar. In: Oller J. W. Jr. (Hrsg.): Issues in language testing research. Rowley, Mass.: Newbury House, S. 3-10.
- Ordinate (2004): SET-10 ® Test description. Validation Summary. Menlo Park, CA: Ordinate. URL: <http://versanttest.de/pdf/ValidationReport.pdf> (Zugriff 18.12.2007).
- Read, J. (2000): Assessing Vocabulary. Cambridge: Cambridge University Press.
- Vollmer, H./Henrici, G./Finkbeiner, C./Grotjahn, R./Schmid-Schönbein, G./Zydati, W. (2001): Lernen und Lehren von Fremdsprachen: Kognition, Affektion, Interaktion. Ein Forschungsüberblick. In: Zeitschrift für Fremdsprachenforschung 12, H. 2, S. 1-145.

Hermann-Günter Hesse / Kerstin Göbel / Nina Jude

17 Interkulturelle Kompetenz

Interkulturelles Lernen wird in den Lehrplänen der Bundesländer als ein Kernelement des Englischunterrichts bezeichnet. Immer wieder werden Zweifel vorgebracht, ob interkulturelles Lernen tatsächlich in nennenswertem Umfang im Englischunterricht stattfindet und inwieweit sich diesbezüglich Lernerfolge bei den Schülern feststellen lassen. Um diesen Fragen nachzugehen, wurde im DESI-Projekt ein Instrument entwickelt, das kulturspezifische interkulturelle Kompetenz bei Schülerinnen und Schülern erfasst. Es kombiniert das Modell der Interkulturellen Sensibilität von Bennett (1993) mit der Methode der kritischen Interaktionssituationen von Thomas (Thomas/Wagner 1999). Darüber hinaus orientierten sich die Inhalte der Testaufgaben an den Lehrplänen der neunten Jahrgangsstufe. Diese beinhalten für den Englischunterricht die Vermittlung kulturell angemessenen Verhaltens, verbunden mit Wissen über die Weltsicht und die Verhaltensorientierungen in englischsprachigen Kulturen. Großbritannien ist ein inhaltlicher Schwerpunkt bis zur neunten Jahrgangsstufe in den verschiedenen Bildungsgängen. Demzufolge beziehen sich die Inhalte des DESI-Instruments zur Erfassung interkultureller Kompetenz auf Interaktionssituationen zwischen Deutschen und Briten. Für die ausführliche Darstellung der Curriculumanalyse im Zusammenhang mit dem Modell von Bennett sei hier auf Göbel und Hesse (2004) verwiesen. Das Instrument berücksichtigt kognitive, affektive und handlungsbezogene Dimensionen, ist also multidimensional angelegt. Anhand ihrer Antworten auf den interkulturellen Test lassen sich die Schülerinnen und Schüler einer von fünf Kompetenzklassen zuordnen, die aus dem zu Grunde gelegten theoretischen Modell abgeleitet wurden. Jede dieser Klassen weist ein spezifisches Profil interkultureller Kompetenz auf. Für die ausführliche Darstellung der Instrumentenentwicklung sei hier auf Hesse/Göbel (2007) verwiesen.

17.1 Interkulturelle Kompetenz bei Schülerinnen und Schülern

Die Lehrpläne der neunten Jahrgangsstufe sehen als Unterrichtsziele des interkulturellen Lernens vor, dass die Lernenden Ethnozentrismus überwinden, Wertschätzung, Interesse und Respekt für kulturelle Verschiedenheit entwickeln, ihre Fähigkeiten zur Anpassung an anderskulturelle Situationen schulen, sich Wissen über kulturelle Werte und Verhaltensnormen sowie über die Dynamik von Kultur und kultureller Identität aneignen. Diese sehr ausgeprägte kulturelle Sensibilität im Sinne ethnorelativer Sichtweisen wird im Modell von Bennett in den Kompetenzklassen „Acceptance“ und „Adaptation“ aufgegriffen. Ihr gegenüber stehen die drei ethnozentrischen Klassen „Denial“ im Sinne der Ignoranz kultureller Unterschiede, „Defense“ im Sinne wertender Ablehnung einer anderen Kultur und „Minimization“ im Sinne des

Herunterspiels kultureller Unterschiede. Die Verteilung der Schülerpopulation auf diese fünf Klassen interkultureller Kompetenz ist in Abbildung 17.1 dargestellt.

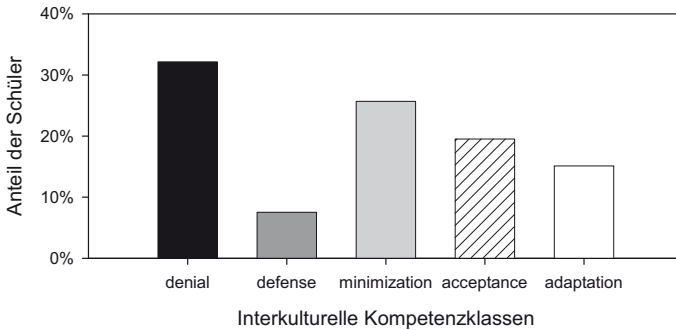


Abbildung 17.1: Verteilung der Klassen interkultureller Kompetenz.

Die Verteilung der IKK-Klassen zeigt, dass etwa ein Drittel der Schülerinnen und Schüler einer Klasse zuzuordnen ist, die sich mit Ignoranz kultureller Unterschiede (Klasse „Denial“) beschreiben lässt. Diese Lerner interpretieren die vorgegebenen kritischen interkulturellen Interaktionssituationen so, als ob kulturelle Unterschiede keine Rolle spielten. Dem stehen nur 8% der Befragten gegenüber, die die englische Kultur direkt ablehnen (Klasse „Defense“). Das Herunterspielen kultureller Unterschiede (Klasse „Minimization“) ist mit 25% der Befragten ebenfalls relativ stark besetzt. Demgegenüber kann die Akzeptanz anderer Kulturen (Klasse „Acceptance“) und die Fähigkeit, sich ihnen anzupassen (Klasse „Adaptation“) als Ausdruck ethnorelativer Orientierungen angesehen werden. Ein beachtlicher Anteil, nämlich 35% der Schülerinnen und Schüler lässt sich diesen Stufen zuordnen. Diesen ethnorelativen Klassen stehen zusammengefasst etwa 40% der Schülerinnen und Schüler in den ethozentrischen Klassen – Defense und Denial – gegenüber. Ein differenzierteres Bild zeigt sich beim Vergleich der verschiedenen Bildungsgänge (vgl. Abbildung 17.2).

Der Anteil an Schülerinnen und Schülern in den ethnozentrischen Kompetenzklassen ist in den Hauptschulen mit 59% und den integrierten Gesamtschulen mit 50% sehr hoch. Deutlich niedriger liegt der Anteil in der Realschule (39%) und im Gymnasium (20%). Entsprechend häufiger finden sich die ethnorelativen Kompetenzklassen in der Realschule (35%) und besonders im Gymnasium wieder, wo über die Hälfte der Schülerinnen und Schüler (55%) Akzeptanz und Anpassung an andere Kulturen in ihren Antworten zum Ausdruck bringen.

Die interkulturellen Kompetenzklassen sind über die Bildungsgänge hinweg ungleich verteilt. Gleichzeitig besteht ein deutlicher Zusammenhang zwischen Bildungsgang, sozioökonomischem Hintergrund der Familie und allgemeinen kognitiven Fähigkeiten der Schülerinnen und Schüler (vgl. Kap. 25). Daraus folgt auch eine Kopplung des familiären sozioökonomischen Hintergrunds mit den Ergebnissen des kognitiven Fähigkeitstests und der Zuordnung zu den interkulturellen Kompetenzklassen.

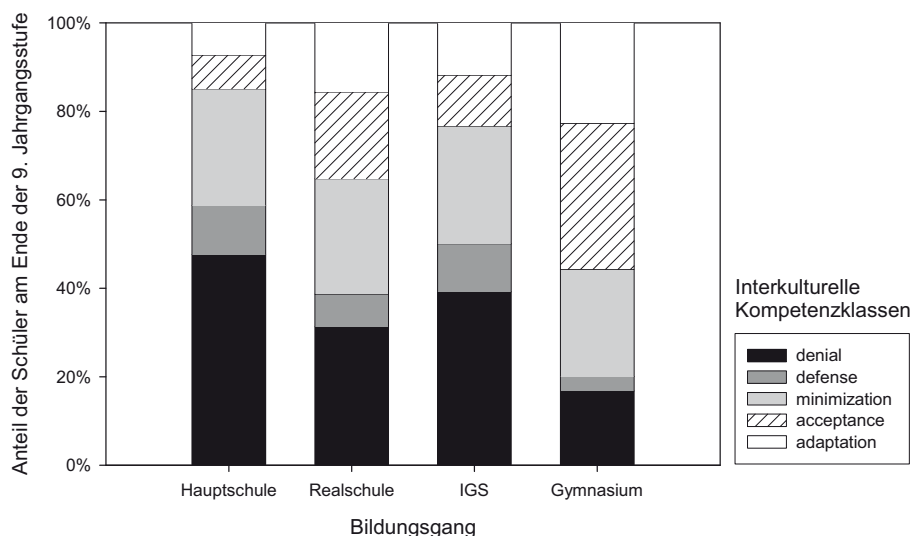


Abbildung 17.2: Prozentuale Verteilung der Schülerinnen und Schüler über die Klassen interkultureller Kompetenz nach Bildungsgängen.

Die Schülerinnen und Schüler, die den Klassen „Denial“ und „Defense“ angehören, stammen eher aus Familien mit niedrigerem sozioökonomischen Hintergrund ($M = 45$), umgekehrt sind die Klassen „Acceptance“ und „Adaptation“ häufiger bei Schülerinnen und Schülern aus Familien mit höherem sozioökonomischen Hintergrund ($M = 52$) anzutreffen; die Schülergruppe der Klasse „Minimization“ liegt mit einem HISEI-Mittelwert von 48 dazwischen. Die Unterschiede sind signifikant, auch in Bezug auf den allgemeinen kognitiven Fähigkeitstest. Analog ergeben sich unterschiedliche Verteilungen der interkulturellen Kompetenzklassen in Bezug auf die Ergebnisse des allgemeinen kognitiven Fähigkeitstests. In der Zugehörigkeit zu den verschiedenen Klassen interkultureller Kompetenz lassen sich weiterhin Geschlechtsunterschiede feststellen (vgl. Abbildung 17.3).

Die weiblichen Schülerinnen sind mehrheitlich den ethnorelativen Klassen „Acceptance“ und „Adaptation“ (zusammen 43%) zuzuordnen, während die männlichen Schüler den ethnozentrischen Klassen „Denial“ und „Defense“ (zusammen 48%) zuzurechnen sind. Der Anteil der männlichen Jugendlichen, die eine „Defense“-Orientierung zeigen, ist mit 12% drei mal so groß wie die der weiblichen Jugendlichen (3%). Der prozentuale Anteil an der Klasse „Minimization“ ist bei beiden Geschlechtern mit 26% gleich groß. Darüber hinaus ist eine Differenzierung der Kompetenzklassen nach dem sprachlichen Hintergrund der Schülerinnen und Schüler interessant (vgl. Abbildung 17.4).

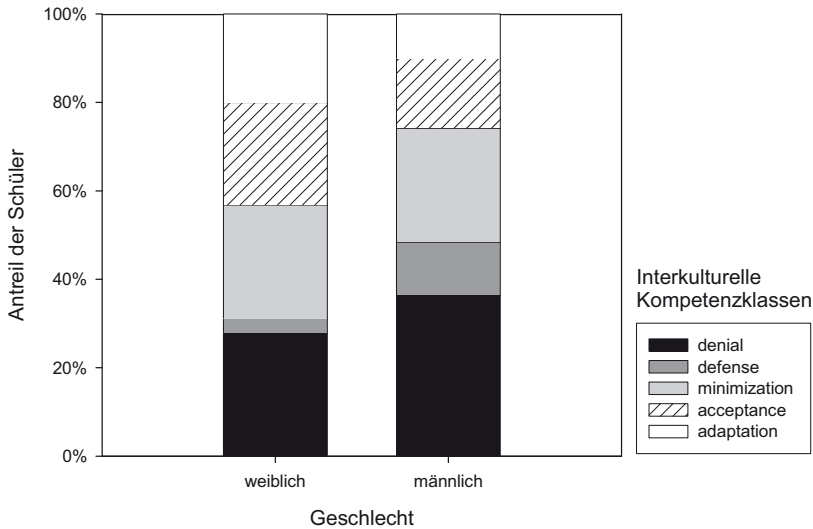


Abbildung 17.3: Prozentuale Verteilung der Schülerinnen und Schüler über die Klassen interkultureller Kompetenz nach Geschlecht.

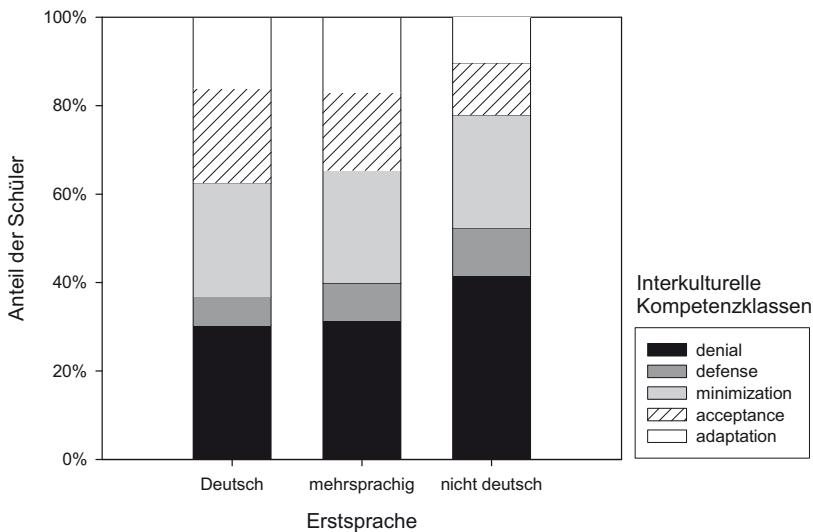


Abbildung 17.4: Prozentuale Verteilung der Schülerinnen und Schüler über die Klassen interkultureller Kompetenz in Bezug auf die Erstsprachen der Schülerinnen und Schüler.

Ein großer Teil der Schülerinnen und Schüler (53%), die eine andere Erstsprache als Deutsch angeben, ist den ethnozentrischen Klassen „Denial“ und „Defense“ zugeordnet. Nur 23% fallen in die beiden ethnorelativen Klassen „Acceptance“ und „Adaptation“. Im Vergleich dazu sind 37% der Schülerinnen und Schüler mit deutscher Erstsprache und 35% der Lernenden die neben Deutsch noch eine oder mehrere weitere Erstsprachen erworben haben („Mehrsprachige“), in diesen beiden Klassen

(zur Definition der drei Erstsprachengruppen vgl. Kapitel 20). Auffallend sind die ähnlichen Verteilungen der Schüler deutscher Erstsprache und der Mehrsprachigen.

Die Ergebnisse zur kulturspezifischen interkulturellen Kompetenz zeigen sich damit ähnlich sensitiv gegenüber den Unterschieden in der sozioökonomischen Herkunft und in den Bildungsgängen sowie den allgemeinen kognitiven Fähigkeiten, wie dies auch bei den Leistungsergebnissen in Deutsch und Englisch zu beobachten ist. Gleichmaßen korrespondiert der Unterschied zwischen weiblichen und männlichen Jugendlichen mit den Leistungsergebnissen im Deutschen und im Englischen. Im Verhältnis zu den Jungen ist ein größerer Anteil der Mädchen den ethnorelativen Klassen zuzuordnen. Der Sachverhalt, dass die Schülerinnen und Schüler nicht deutscher Erstsprache zum großen Teil den ethnozentrischen interkulturellen Klassen zugeordnet wurden, entspricht darüber hinaus den Erwartungen, da sie überwiegend aus Familien mit niedrigem sozioökonomischen Status stammen.

17.2 Modellorientierte Analyse

Wir stellen uns zunächst die Frage, ob die Ausprägungen der interkulturellen Kompetenz von individuellen und damit kaum oder gar nicht von beeinflussbaren Faktoren und außerschulischen Lerngelegenheiten abhängen. In welcher Weise sich schulischer Unterricht auf die interkulturelle Kompetenz auswirkt, ist in Kapitel 33 dargestellt. Wegen der Vermischung der potenziellen Erklärungsvariablen wurde ein Modell der Zweiebenenanalyse gewählt, das in der Lage ist, die Effekte der konfundierten Faktoren zu kontrollieren und über ihren Einfluss hinaus die Wirkungen von Lerngelegenheiten zu erfassen (vgl. Kapitel 3). Die Ergebnisvariable erfasst die Wahrscheinlichkeiten der Zuordnung eines Schülers zu einer der interkulturellen Kompetenzklassen. Aus Gründen der Darstellungsvereinfachung werden im Folgenden die Klassen „Denial“ und „Defense“ zur neuen Klasse „Denial/Defense“ (abgekürzt „DD“), die Klassen „Acceptance“ und „Adaptation“ zur neuen Klasse „Acceptance/Adaptation“ (abgekürzt „AA“), zusammengefasst, während die Klasse „Minimization“ (abgekürzt „M“), gesondert bestehen bleiben soll (vgl. Hesse/Göbel 2007). Zur Analyse wurde ein dreikategoriales multinomiales Zweiebenenmodell formuliert bei dem drei interkulturelle Kompetenzklassen als „abhängige Variable“ definiert sind (Raudenbusch u.a. 2001). 3900 Schüler aus 421 Klassen erhielten die entsprechenden Fragebögen und gingen somit in die folgenden Analysen ein.

Auf der Individualebene wurden die folgenden Variablen in die Mehrebenenanalysen aufgenommen:

- Der sozio-ökonomischen Status (HISEI), die kognitiven Fähigkeiten (KFT), die Erstsprache und das Geschlecht.
- Die in einem Gesamtwert zusammengefassten Leistungen in den DESI-Englischtests und in Soziopragmatik, die individuellen Selbstwirksamkeitsüberzeugungen und das fachbezogene Fähigkeitskonzept.
- Die Nutzung von außerschulischen Lerngelegenheiten durch Urlaub im englischsprachigen Ausland, Austausch, Sprachreisen, Kontakten mit englischsprachigen Personen sowie die Wertschätzung des Englischen in der Familie.

Auf der Klassenebene wurden die Effekte der Faktoren Bildungsgang, durchschnittlicher sozioökonomischer Status, durchschnittliches kognitives Fähigkeitsniveau in der Klasse, Anteil der Schülerinnen und Schüler nichtdeutscher Erstsprache, Anteil der mehrsprachigen Schüler und Anteil der Mädchen in der Schulklasse geprüft. Keiner der diesbezüglichen Effekte stellte sich auf Klassenebene als statistisch bedeutsam heraus. Weder der mittlere sozioökonomische Status, das durchschnittliche Niveau kognitiver Grundfähigkeiten, das Verhältnis Schülerinnen zu Schüler, noch der Anteil von Schülerinnen und Schülern, die eine andere Erstsprache als Deutsch angeben, wirken sich auf die interkulturelle Kompetenz, so wie sie in DESI gemessen wurde, aus. Auch die Zugehörigkeit zu einem der Bildungsgänge spielt keine Rolle. Auf Individualebene finden sich hingegen signifikante Differenzen.

In Abbildung 17.5 sind die Wahrscheinlichkeiten abgebildet, mit denen eine Schülerin oder ein Schüler in Abhängigkeit von den individuellen Merkmalen, einer der drei interkulturellen Kompetenzklassen „Denial/Defense“, „Minimization“ oder „Acceptance/Adaptation“ zugeordnet wurde.

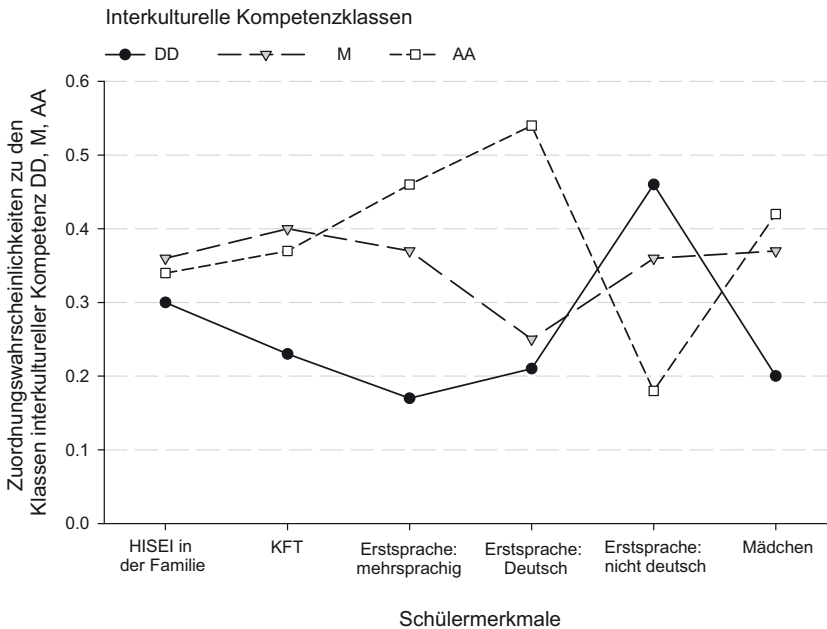


Abbildung 17.5: Zuordnungswahrscheinlichkeiten zu den Klassen interkultureller Kompetenz in Abhängigkeit der Schülermerkmale auf Individualebene – HISEI, KFT, Erstsprache und Geschlecht.

Bei simultaner Berücksichtigung der Kontextfaktoren treten die statistischen Effekte des sozioökonomischen Status auf der individuellen Ebene in den Hintergrund. Schülerinnen und Schüler mit einem KFT-Wert, der eine Standardabweichung über dem Durchschnitt liegt, werden mit einer Wahrscheinlichkeit von .40 „Acceptance/Adaptation“ zugeordnet, von .37 „Minimization“ und von .23 der ethnozentrischen Klasse „Denial/Defense“ zugeordnet. Signifikant sind die Unterschiede zwischen den

beiden Klassen „Acceptance/Adaptation“, „Minimization“ und der Klasse „Denial/Defense“. Das bedeutet, dass die Ignoranz oder die Abwertung der englischen Kultur mit niedrigeren kognitiven Grundfähigkeiten korrespondiert.

Simultan mehrsprachig aufgewachsene Schülerinnen und Schüler bearbeiten die interkulturellen Testaufgaben mit einer mittleren Wahrscheinlichkeit von .46 in einer Weise, die als „Acceptance“ oder „Adaptation“ zu charakterisieren ist. Im Vergleich dazu liegt die Wahrscheinlichkeit für „Minimization“ bei .37 und für „Denial“ oder „Defense“ bei .17. Der Kontrast DD-AA ist signifikant. Interessanterweise besteht für die monolingual deutschsprachige Schülerschaft eine hohe Wahrscheinlichkeit (.54) der Klasse „Acceptance/Adaptation“ und andererseits eine niedrigere Wahrscheinlichkeit den Klassen „Minimization“ (.25) und „Denial/Defense“ (.21) zugeordnet zu werden. Die Kontraste zwischen DD-AA und M-AA sind signifikant. Die für die Mehrsprachigen und deutsch Erstsprachigen beobachtete Abfolge der Zuordnungswahrscheinlichkeiten dreht sich für die nicht deutsch Erstsprachigen um: Sie gehören mit einer Wahrscheinlichkeit von .46 der ethnozentrischen Klasse „Denial/Defense“ an, mit einer geringeren Wahrscheinlichkeit (.18) zur ethnorelativen Klasse „Acceptance/Adaptation“ sowie mit einer Wahrscheinlichkeit von .33 zur Klasse „Minimization“. Die Unterschiede zwischen DD-AA und M-AA sind signifikant.

Mehrsprachigkeit und eine von Deutsch verschiedene Erstsprache indizieren in der überwiegenden Zahl der Fälle einen Migrationshintergrund (vgl. Kapitel 20). Die Analysen zeigen, dass die Schülerschaft mit einer nicht deutschen Erstsprache mehr Schwierigkeiten zu besitzen scheint, eine ethnorelative Sicht im interkulturellen Test zu äußern als die anderen Schülergruppen. Im Gegensatz dazu neigen offenbar Mehrsprachige selten zu ausgesprochen ethnozentrischen Äußerungen. Dies unterstützt die These, dass Mehrsprachigkeit einen Faktor darstellt, der zur Reduktion der Ignoranz oder der Ablehnung anderer Kulturen beiträgt (vgl. Kapitel 20).

Mädchen werden mit einer höheren Wahrscheinlichkeit den ethnorelativen Klassen „Acceptance/Adaptation“ (.42) bzw. der Klasse „Minimization“ (.37) und seltener der Kompetenzklasse „Denial/Defense“ (.21) zugeordnet, wobei die Kontraste DD-M und DD-AA signifikant sind. Dieses Ergebnis spricht für eine hohe Sensibilität von Schülerinnen gegenüber interkulturellen Themen im Englischunterricht.

Bei den bisher dargestellten Merkmalen handelt es sich um unveränderbare Eingangsvoraussetzungen der Schülerschaft. Es soll nun untersucht werden, inwieweit die Gesamtestleistung im Englischen und die Ergebnisse im Bereich der Soziopragmatik im Englischen mit der Zuordnung zu interkulturellen Kompetenzklassen in Zusammenhang gebracht werden können. Die Soziopragmatik ist ein Ausdruck für die Fähigkeit, sich im Sprachgebrauch angemessen in einer kommunikativen Situation verhalten zu können, was im Bereich der interkulturellen Kommunikation eine herausragende Rolle spielt (vgl. auch Kapitel 20). Es liegt nahe, einen Zusammenhang mit interkultureller Kompetenz zu vermuten. Weiterhin soll geprüft werden, inwieweit sich die Selbstwirksamkeitsüberzeugung und die Einschätzung des fachbezogenen Fähigkeitskonzepts der Lernenden auf die interkulturelle Kompetenz auswirken (vgl. Abbildung 17.6).

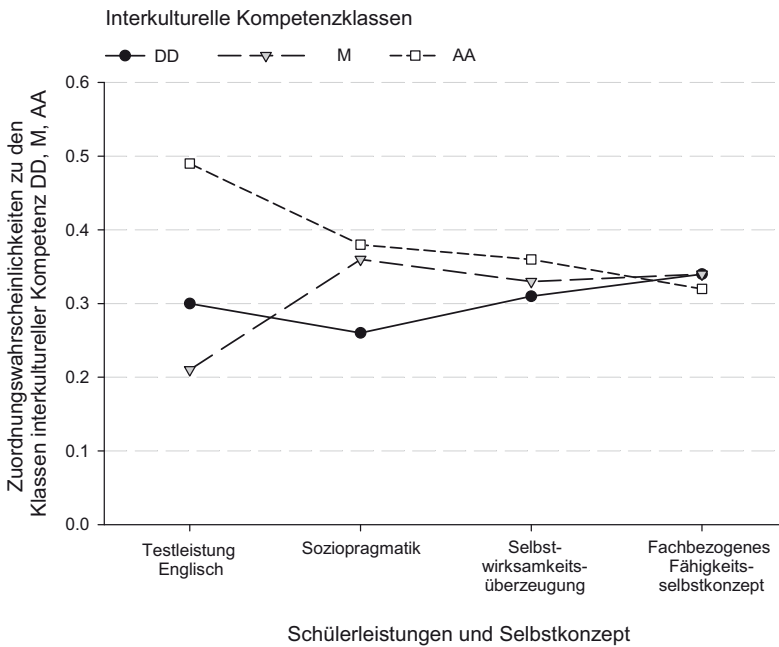


Abbildung 17.6: Zuordnungswahrscheinlichkeiten zu den Klassen interkultureller Kompetenz in Abhängigkeit der Schülermerkmale auf Individualebene – Testleistungen und Selbstkonzept.

Wie in Abb. 17.6 erkennbar, weisen lediglich die Leistungen im Gesamttest für Englisch und die Leistungen im Soziopragmatik-Test statistisch signifikante Effekte auf die Zuordnung zu den Klassen interkultureller Kompetenz auf. Die Selbstwirksamkeitsüberzeugung und das fachbezogene Fähigkeits-selbstkonzept der Schülerinnen und Schüler stehen in keinem Zusammenhang mit der Zuordnung zu den interkulturellen Kompetenzklassen

Fähigkeitsbezogene Einschätzungen der Schülerinnen und Schüler in Bezug auf das Englischlernen haben demnach nur eine schwach differenzierende Wirkung auf die Zugehörigkeit zu einer interkulturellen Kompetenzklasse. Die Zuordnungswahrscheinlichkeiten auf Grund von Selbstwirksamkeitsüberzeugungen und dem fachbezogenen Fähigkeitskonzept im Englischen variieren nicht signifikant.

Die Lernenden mit einer hohen Gesamttestleistung in Englisch haben eine hohe Wahrscheinlichkeit (.49), den ethnorelativen Klassen „Acceptance/Adaptation“ zugeordnet zu werden. Die Zuordnung zu „Minimization“ (.30) und „Denial/Defense“ (.21) ist für diese Gruppe weniger wahrscheinlich. Die Kontraste DD-AA und M-AA sind signifikant. Hohe Testwerte im Soziopragmatik-Test gehen ebenfalls mit einer höheren Wahrscheinlichkeit (.38) einher, den Klassen „Acceptance/Adaptation“ zugeordnet zu werden. Im Vergleich hierzu ist die Wahrscheinlichkeit für eine Zuordnung dieser Schülergruppe zu „Denial/Defense“ (.26) deutlich geringer. Der Unterschied DD-AA ist signifikant.

Es liegt die Vermutung nahe, dass neben den individuellen Merkmalen wie den kognitiven Fähigkeiten oder der Erstsprache weitere außerschulische Einflussfaktoren wirksam sind. Daher wurden die Zusammenhänge zwischen sprachspezifischen, außerschulischen Lerngelegenheiten in der Familie und der interkulturellen Kompetenz von Schülerinnen und Schülern analysiert. Wie Abbildung 17.7 zeigt, scheinen die Wirkungen der außerschulischen Lerngelegenheiten jedoch nicht so durchschlagend zu sein, als dass es – bei Kontrolle aller anderen genannten Faktoren – zu signifikanten Unterschieden kommen würde.

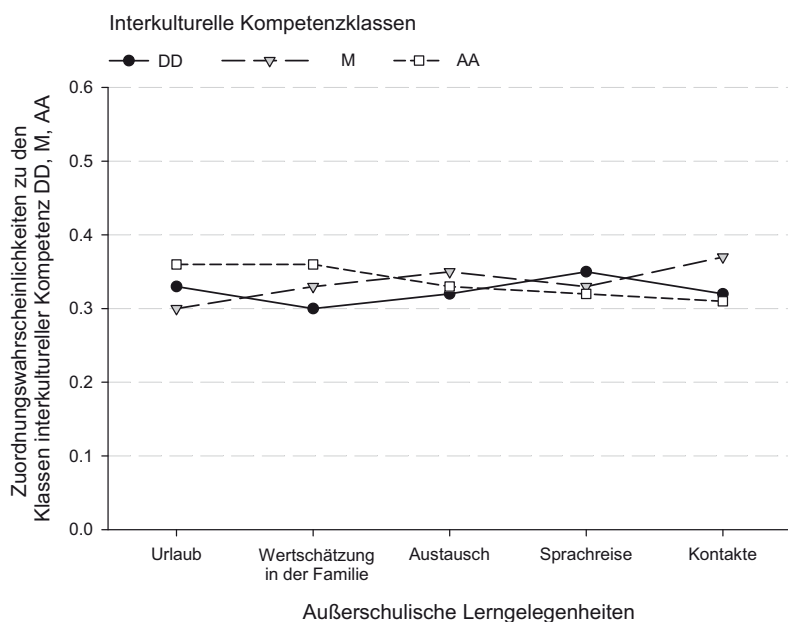


Abbildung 17.7: Zuordnungswahrscheinlichkeiten zu den Klassen interkultureller Kompetenz in Abhängigkeit der Schülermerkmale auf Individualebene – außerschulische Lerngelegenheiten.

Entgegen der berechtigten Annahme, dass interkulturelle Kompetenz durch Aufenthalte im englischsprachigen Ausland wie Sprachreisen oder Schüleraustausch gefördert wird, lassen sich hier keine signifikanten Unterschiede in den Zuordnungswahrscheinlichkeiten feststellen. Lediglich die Häufigkeit des Urlaubs im englischsprachigen Ausland erhöht die Wahrscheinlichkeit, den Klassen „Acceptance/Adaptation“ im Vergleich zu „Minimization“ zugeordnet zu werden. Dieser Kontrast ist signifikant. Die Wertschätzung der Sprache Englisch in der Familie und private Kontakte zu Personen im englischsprachigen Ausland können die Zuordnung zu den Klassen interkultureller Kompetenz nicht vorhersagen.

17.3 Diskussion

Zusammenfassend lässt sich feststellen, dass die individuelle interkulturelle Kompetenz in hohem Maße mit individuellen Voraussetzungen und persönlichen Merkmalen zusammenhängt. Dazu gehören die allgemeinen kognitiven Fähigkeiten, die Erstsprache und das Geschlecht. Mädchen werden eher den ethnorelativen Klassen zugeordnet als Jungen, ebenso hängen die kognitiven Eingangsvoraussetzungen positiv mit ethnorelativen Haltungen (Acceptance/Adaptation) zusammen. Dass weiterhin Schülerinnen und Schüler nicht deutscher Erstsprache eine höhere Wahrscheinlichkeit haben, den ethnozentrischen Klassen zugeordnet zu werden, hängt vermutlich mit dem insgesamt eher niedrigen sozioökonomischen Status dieser Gruppe zusammen. Die Wirkung der sozioökonomischen Herkunft und der Bildungsgänge auf die Zuordnung zu den interkulturellen Klassen tritt bei gleichzeitiger Kontrolle der Kontextfaktoren zurück. Ein deutlicher Zusammenhang besteht hingegen, auch nach der Kontrolle der individuellen Merkmale auf Klassenebene, zwischen der interkulturellen Kompetenz und der Gesamttestleistung im Englischen sowie dem Soziopragmatik-Test. Hieraus lässt sich der Schluss ziehen, dass die Schülerinnen und Schüler, die gute Leistungen in den Englischtests erzielen, spezifische interkulturelle Begegnungssituationen besser als solche identifizieren können. Damit einhergehend erfüllen sie mit einer höheren Wahrscheinlichkeit die Erwartungen der Lehrpläne in Bezug auf interkulturelle Kompetenz, nämlich die erkannten kulturellen Unterschiede zu akzeptieren und sich in einem gewissen Umfang an die anderskulturelle Umwelt anpassen zu können. Das Wissen über sozial adäquate Verhaltensweisen in einer anderen Kultur – im DESI-Test Soziopragmatik erfasst – scheint mit dem DESI-Instrument Interkulturelle Kompetenz in besonderer Weise zusammenzuhängen. Außerschulische Lerngelegenheiten, so wie sie von den Schülerinnen und Schülern berichtet werden, scheinen demgegenüber bis auf Urlaubsreisen in ein englischsprachiges Land kaum eine Wirkung auf die in DESI gemessene interkulturelle Kompetenz zu haben. Die Analysen legen die Annahme nahe, dass eine unsystematische Förderung interkulturellen Lernens bei Schülerinnen und Schülern dieser Altersgruppe nur bedingt Effekte auf die Lernergebnisse hat. Vor dem Hintergrund des starken Einflusses von Individualmerkmalen einerseits und der geringen Bedeutung der erhobenen außerschulischen Förderung andererseits, stellt sich die Frage, inwieweit der Unterricht die individuellen Voraussetzungen kompensieren kann. In welcher Weise eine systematische Förderung im Englischunterricht zur interkulturellen Kompetenz der Lernenden beiträgt, wird in Kapitel 33 behandelt.

Literatur

- Bennett, M.J. (1993): Towards ethnorelativism: A developmental model of intercultural sensitivity. In: Paige, M. R. (Hrsg.): Education for the intercultural experience. Yarmouth, MN: Intercultural Press S. 21-71.
- Göbel, K./Hesse, H.-G. (2004): Vermittlung interkultureller Kompetenz im Englischunterricht – eine curriculare Perspektive. In: Zeitschrift für Pädagogik 50, H. 6, S. 818-834.

- Hesse, H.-G./Göbel, K. (2007): Interkulturelle Kompetenz. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band I. Weinheim: Beltz, S. 256-272.
- Raudenbush, S./Bryk, A./Cheong, Y. F./Congdon, R. (2001): *HLM 5. Hierarchical linear and non-linear modeling*. Lincolnwood, IL: Scientific Software International.
- Thomas, A./Wagner, K.H. (1999): Von der Fremdheitserfahrung zum interkulturellen Verstehen. Kulturpsychologische Grundlagen für den Einsatz interkultureller Trainingsprogramme. In: Praxis des neusprachlichen Unterrichts 46, S. 227-236.

*Nina Jude / Eckhard Klieme / Wolfgang Eichler / Rainer H. Lehmann /
Günter Nold / Konrad Schröder / Günther Thomé / Heiner Willenberg*

18 Strukturen sprachlicher Kompetenzen

Sprachliche Kompetenzen sind ein wesentlicher Bestandteil von Bildung und werden durch schulischen Unterricht in unterschiedlicher Weise beeinflusst, so dass sie in empirischen Studien zu Schülerleistungen immer stärkere Beachtung finden (Bos u.a. 2003; Deutsches PISA Konsortium 2001). Neben der Feststellung von Ausprägungen der sprachlichen Kompetenzen, also der Einordnung von Schülerinnen und Schülern auf Kompetenzniveaus, stellt sich vermehrt auch die Frage nach der Differenzierung von Sprachkompetenzen, also danach, welche spezifischen sprachlichen Phänomene wie gut beherrscht werden. Aus sprachwissenschaftlicher und fachdidaktischer Sicht ist danach zu fragen, welche Zusammenhänge zwischen einzelnen sprachlichen Teilbereichen bestehen: Gibt es beispielsweise systematische Gemeinsamkeiten im Lesen und Schreiben, so dass ein übergeordneter, schriftsprachlicher Kompetenzfaktor angenommen werden kann? Oder sind die Zusammenhänge zwischen Leseverstehen und Hörverstehen so hoch, dass sich eine rezeptive Sprachkomponente damit beschreiben lässt? Wie lassen sich also Strukturen von Sprachkompetenz abbilden, wenn dazu Daten aus Schulleistungstudien analysiert werden?

Aus entwicklungspsychologischer Perspektive stellt der Spracherwerb einen sehr komplexen Prozess dar, der sich über einen langen Zeitraum erstreckt. Dabei werden nicht alle Teilfähigkeiten gleichzeitig und gleich gut erlernt (Szagun 2000; Haberzettl/Wegener 2003). Hörverstehen und Sprechen zählen zu den früh in der menschlichen Entwicklung ausgebildeten Kompetenzen. Der Erwerb von schriftsprachlichen Fähigkeiten beginnt meist mit der Einschulung. Auch die institutionelle Vermittlung weiterer spezifischer Teilkompetenzen wie bspw. von Wortschatz und Grammatik erfolgt dann in schulischen Lehr-Lernkontexten (Detering 2000; Eckerth 2000). Es wird angenommen, dass sich mit zunehmender gesamtsprachlicher Kompetenz die Ausprägungen der unterschiedlichen Facetten einander annähern, d.h. dass die schriftsprachlichen Leistungen in der Erstsprache den auditiven entsprechen und ihre Entwicklung mit dem Eintritt ins Erwachsenenalter größtenteils abgeschlossen ist (Oerter/Montada 2002). Die Strukturen von Sprachkompetenz bei Erwachsenen würden sich in diesem Fall in einem homogenen, globalen Faktor abbilden lassen.

Entsprechende Modelle finden sich in linguistischen Ansätzen der 1970er Jahre, die von einer singulären übergreifenden Sprachkompetenz ausgehen und alle spezifischen sprachlichen Fähigkeiten hierauf beziehen (Oller 1976; Kunnan 1992). Inzwischen legen jedoch aktuelle Forschungsergebnisse nahe, dass die Struktur von Sprachkompetenz als mehrdimensional betrachtet werden muss, und in Abhängigkeit von Alter oder Sprachhintergrund der untersuchten Personengruppen durchaus variieren kann (North 2000). Neue kommunikative Ansätze in der Leistungserhebung,

die stärker offene und mündliche Testformate verwenden, unterstützen die facettenreiche Messung von Sprachkompetenz ebenso wie statistische Analysemethoden, die die Überprüfung solcher mehrdimensionaler Modelle ermöglichen (McNamara 1995; Kuhlemeier/van den Bergh/Rijlaarsdam 2002). So können methodische Artefakte reduziert und differenzierte Modelle von Sprachkompetenz konfirmatorisch überprüft werden (Gräfe-Bentzien 2000).

Die Mehrzahl der psycholinguistischen Modelle in diesem Forschungsfeld postulieren vier Basiskomponenten – Lesen, Schreiben, Hörverstehen und Sprachproduktion – die sich wiederum den Prozessen der Sprachproduktion und -rezeption bzw. den Modalitäten der schriftlichen und auditiven Sprachverwendung zuordnen lassen. Ein differenziertes Modell der Struktur von Sprachkompetenz mit exemplarisch ausgewählten sprachlichen Teilbereichen ist in Tabelle 18.1 dargestellt.

Tabelle 18.1: Prozess-Modalitätsmodell mit differenzierten Facetten sprachlicher Kompetenz.

Modalität Prozess	Auditiv	Schriftsprachlich
Produktion	Sprechen	Schreiben
	Grammatik Wortschatz Aussprache	Grammatik Wortschatz Rechtschreiben
Reflektion Rezeption	Hören	Lesen
	Grammatik Wortschatz Wahrnehmungstempo	Grammatik Wortschatz Lesegeschwindigkeit

Aus Perspektive der Lehr-Lernforschung interessiert, wie diese Teildimensionen zusammenhängen und welche Struktur der Sprachkompetenz sie, ggf. in Faktoren höherer Ordnung, bilden. So könnten sich die Leistungen im Lesen und Hörverstehen auf die gemeinsame Dimension „Rezeption“ zurückführen lassen (Rost/Hartmann 1992). Alternativ könnte der Zusammenhang zwischen Sprechen und Hören stärker ausgeprägt sein und sich somit die Dimension „auditive Sprachkompetenz“ abbilden lassen. Solche Zusammenhänge können Rückschlüsse auf gemeinsame kognitive Vorgänge oder spezifische Lehr-Lernansätze ermöglichen: Ein stark kommunikativ ausgerichteter Sprachunterricht könnte dazu führen, dass sich die auditiven sprachlichen Teilkompetenzen Sprechen und Hörverstehen ähnlich stark ausprägen. Entsprechend könnte ein eher schriftbasierter Unterricht die schriftsprachlichen Subkomponenten parallel fördern.

18.1 Theoretische Modelle sprachlicher Kompetenzen

Die Leistungen im Deutschen und Englischen wurden in DESI in jeweils sechs unterschiedlichen Bereichen erfasst (vgl. Kapitel 2), dabei kamen verschiedene Testformate zum Einsatz, wobei der Schwerpunkt auf der schriftgestützten Erfassung von Sprachkompetenz lag (Beck/Klieme 2007). Der Test zur mündlichen Sprachproduktion im Englischen (Set-10, vgl. Kapitel 16) wurde lediglich von einer

kleinen Stichprobe von Schülern bearbeitet und wird in die Gesamtanalysen nicht einbezogen.

Die Auswahl der sprachlichen Kompetenzen wurde durch die Zielsetzung der DESI-Studie bestimmt. Anhand der erhobenen Daten ist es möglich, eine Vielzahl unterschiedlicher, theoretisch fundierter Modelle zu analysieren. Besonders interessant für DESI ist dabei die Frage, wie differenziert die Kompetenzstruktur auf Individualebene abgebildet werden kann. Je besser dies gelingt, umso deutlicher stellen sich die unterschiedlichen Testkonstrukte dar, welche die vielfältigen sprachlichen Teilbereiche in DESI abbilden sollen. Im Folgenden werden die in Abb. 18.1 dargestellten Modelle vergleichend analysiert:

- Modelle, die einen *Gesamtfaktor* der Sprachkompetenz annehmen (G-Faktor-Modell),
- Modelle, die eine *mehrdimensionale* Differenzierung korrelierender sprachlicher Teilkompetenzen postulieren und
- Modelle, die spezifische *produktive und rezeptive* Faktoren annehmen.

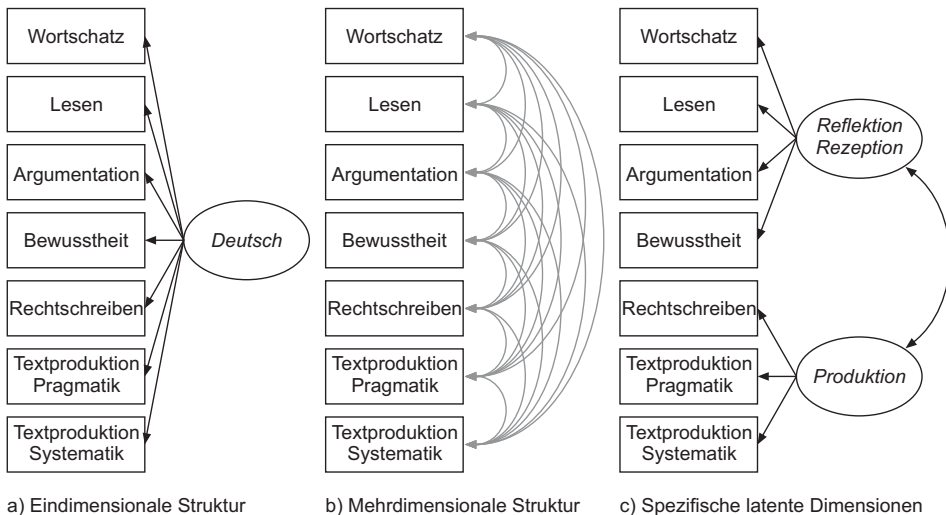


Abbildung 18.1. Unterschiedliche Modelle der Struktur von Sprachkompetenz am Beispiel Deutsch.

Die Analyse dieser Modelle von Sprachkompetenz erfolgt mittels Strukturmodellierungen, die auf Kovariationen der unterschiedlichen Teilttests miteinander basieren (Kunnan 1998; Muthén 2004). Diese Kovariationen können sich in ihrem Ausmaß zwischen der Individualebene und der Klassenebene unterscheiden, d.h. innerhalb einer Klasse können die Leistungen der Schüler homogener oder heterogener sein als die von Schülern in einer anderen Klasse (Van den Noortgate/Opendakker/Onghena 2005). Um dieses Phänomen auch statistisch zu berücksichtigen wird daher die Struktur von Sprachkompetenz mittels Mehrebenen-Faktorenanalysen auf zwei Ebenen modelliert: Auf Individualebene und auf Klassenebene (Hox 1993, 2002; vgl. auch Kapitel 3). Um zu beurteilen, welches theoretische Modell den Daten der

DESI-Studie am besten gerecht wird, werden die drei oben genannten Modelle (G-Faktor-Modell, mehrdimensionales Modell und hierarchisches Modell) jeweils gegenübergestellt und verglichen (Jöreskog 1993). Die Vergleichskriterien sind dabei die bei statistischen Modellvergleichen gängigen Modellgütekriterien (Tanaka 1993; Schermelleh-Engel/Moosbrugger 2002).

18.2 Allgemeines Modell der Deutschkompetenz

Die DESI-Testmodule erfassen mit unterschiedlichen Testformaten verschiedene Aspekte der Deutschkompetenz. Nicht nur dieses methodische Vorgehen, sondern auch bereits vorhandene empirische Befunde lassen erwarten, dass sich eine differenzierte Kompetenzstruktur abbilden lässt. In der Tat weist die Modellierung eines rein eindimensionalen Modells (G-Faktor), das auf Individual- und auf Klassenebene nur einen gemeinsamen latenten Faktor als Ursache der Leistungsvariationen in den Teilkompetenzen annimmt, lediglich einen befriedigenden Fit auf. Besonders auf Klassenebene weisen die Ergebnisse darauf hin, dass sich nicht alle Varianz der Testmodule durch eine gemeinsame Dimension erklären lässt. Eine Ursache hierfür könnte die testmethodische Umsetzung sein: Die Testmodule Rechtschreiben und Textproduktion stellen Testformate dar, in denen die (freie) Produktion geschriebener Sprache gefordert wird, und unterscheiden sich schon dadurch von den eher geschlossenen Formaten der vier anderen Testmodule. Der Vergleich der drei oben beschriebenen, theoretischen Modelle von Sprachkompetenz weist auf eine Struktur hin, die der Datenlage angemessener erscheint: Das Modell, das auf Individualebene eine mehrdimensionale Struktur sowie auf Klassenebene die hierarchischen Faktoren Sprachrezeption und Sprachproduktion annimmt, weist eine gute Passung auf (vgl. Abbildung 18.2).

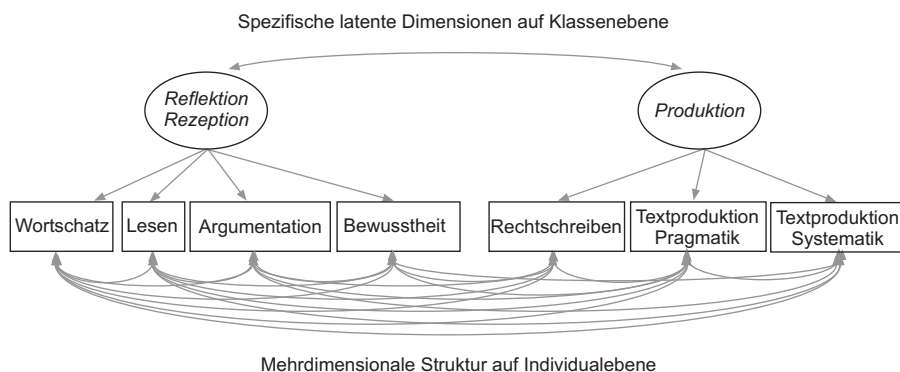


Abbildung 18.2: Mehrebenen-Strukturmodell der Deutschkompetenz in DESI. Modellpassung: $CFI=.99$, $RMSEA=.02$; $SRMR_{within}=.00$, $SRMR_{between}=.01$.

Die Kompetenzstruktur auf Individualebene stellt sich durchaus differenziert dar und verdeutlicht die Heterogenität der sprachlichen Fähigkeiten. Diese korrelieren zwar untereinander, jedoch nicht so hoch, dass von einer eindimensionalen Struktur ge-

sprochen werden kann. Die DESI-Testverfahren erfassen also durchaus differenzierbare Teildimensionen von Sprachkompetenz. Die Korrelationen auf Individualebene sind in Tabelle 18.2 dargestellt.

Tabelle 18.2: Korrelationen der Testmodule Deutsch auf Individualebene.

	Lesen	RS	KA	TP2	TP3	WS
Rechtschreiben (RS)	.15					
Argumentation (KA)	.28	.18				
Textproduktion Pragmatik (TP 2)	.25	.22	.29			
Textproduktion Systematik (TP 3)	.17	.36	.17	.37		
Wortschatz (WS)	.35	.18	.26	.20	.17	
Bewusstheit (BD)	.34	.29	.29	.31	.29	.36

Die Korrelationen der sprachlichen Teilkompetenzen auf Individualebene variieren deutlich: Zwischen Rechtschreiben und Lesen betragen sie lediglich $r = .15$, zwischen Wortschatz und Sprachbewusstheit hingegen $r = .36$, ebenso zwischen Rechtschreiben und Textproduktion Systematik. Die beiden letzteren Teilmodule lassen sich auf Klassenebene auf die gemeinsame Dimension *Produktion* zurückführen. Auf Klassenebene kann darüber hinaus eine zweite differenzierbare latente Dimension, *Rezeption und Reflektion*, abgebildet werden. Diese Dimension bezieht die Indikatoren des Lesens, der Sprachbewusstheit, der Argumentation und des Wortschatzes ein. Die latenten Dimensionen auf Klassenebene korrelieren zu $r = .98$ miteinander. Trotz dieser sehr hohen Korrelation erreicht das Modell mit einer eindimensionalen Struktur auf Klassenebene eine schlechtere Anpassung an die Daten.

18.3 Allgemeines Modell der Englischkompetenz

Die Indikatoren für die Sprachkompetenz im Englischen beinhalten im Unterschied zu den Deutschtests ein Testmodul zum Hörverstehen sowie C-Tests als Indikatoren für die produktiven Fähigkeiten im Englischen. Zur Analyse der Struktur der Englischkompetenz wurden entsprechend den Überlegungen zur Deutschkompetenz drei Modelle miteinander verglichen (vgl. Abbildung 18.3).

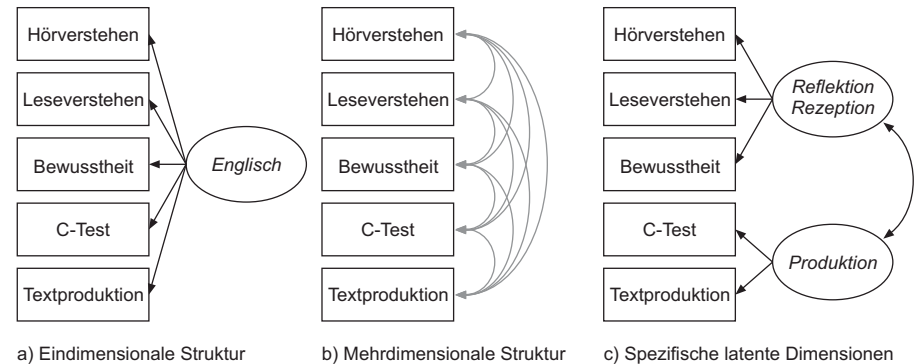


Abbildung 18.3: Unterschiedliche Modelle der Struktur der Englischkompetenz.

Auch hier zeigt sich analog zur Kompetenz im Deutschen, dass nicht pauschal von einer Eindimensionalität der Fremdsprachenkompetenz ausgegangen werden kann. Im Vergleich der verschiedenen Modelle für die Englischkompetenz der DESI Schülerinnen und Schüler lässt sich das in **Abbildung 18.4** dargestellte Mehrebenenmodell abbilden: Auf Individualebene besteht eine mehrdimensionale Struktur mit korrelierenden, differenzierten Kompetenzen, auf Klassenebene hingegen können die latenten Dimensionen *Sprachreflektion/ Sprachrezeption* und *Produktion* gebildet werden (vgl. **Abbildung 18.4**).

Spezifische latente Dimensionen auf Klassenebene

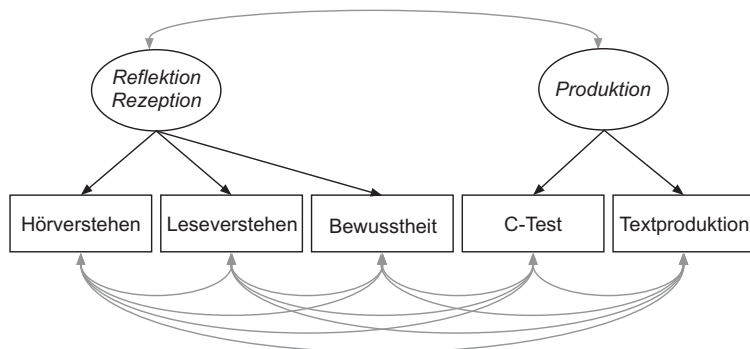


Abbildung 18.4: Mehrebenen-Strukturmodell der Englischkompetenz in DESI. Modellpassung: $\chi^2/df = 2.8$; $p = .02$; $CFI = .99$, $RMSEA = .01$; $SRMR_{within} = .00$, $SRMR_{between} = .00$.

Die Zusammenhänge zwischen den sprachlichen Teilbereichen im Englischen lassen sich parallel zur Struktur im Deutschen beschreiben: Auf Individualebene zeigt sich ein differenziertes, mehrdimensionales Bild, auf Klassenebene stellen sich die latenten Dimensionen der Sprachproduktion sowie der Reflektion-Rezeption dar, letztere korrelieren zu $r = .99$. Die Korrelationen der Teilmodule auf Individualebene sind in **Tabelle 18.3** dargestellt.

Tabelle 18.3: Zusammenhänge der Indikatoren der Englischkompetenz auf Individualebene

	Lesen	HV	CT	BE
Hörverstehen (HV)	.43			
C-Test	.49	.39		
Bewusstheit (BE)	.35	.30	.41	
Textproduktion (ST)	.42	.28	.52	.35

Die Korrelationsmuster auf Individualebene zeigen etwas höhere Zusammenhänge zwischen den Teilmodulen im Englischen, als dies bei den Deutschtests zu beobachten ist. Die höchste gemeinsame Variation findet sich für die Formate Textproduktion und C-Test. Hier werden ähnliche Kompetenzen, nämlich das (freie) Schreiben von Text unter Rückgriff auf Vokabular und Rechtschreiben, gefordert. Die insgesamt höheren Korrelationen könnten sich dadurch erklären lassen, dass der insti-

tutionelle Erwerbsprozess der Fremdsprache alle sprachlichen Bereiche stärker simultan integriert als dies bei der Erstsprache der Fall ist. Auch die Aufgaben zum Lese- und Hörverstehen sowie zur Sprachbewusstheit, die auf Klassenebene sogar einen gemeinsamen Faktor bilden, weisen auf Individualebene zwar hohe gemeinsame Varianz auf, die Annahme einer gemeinsamen latenten Dimension lässt sich jedoch nicht aufrecht erhalten. In der durchaus geringeren Korrelation des Moduls Textproduktion mit dem Hörverstehenstest zeigt sich auch auf Individualebene, dass diese Testverfahren durchaus differenzierbare sprachliche Bereiche ansprechen. Diese lassen sich auch auf Klassenebene in der Differenzierung zwischen Rezeption und Produktion wiederfinden.

18.4 Strukturelle Zusammenhänge zwischen dem Deutschen und dem Englischen

Nachdem die Struktur der Sprachkompetenz im Deutschen und im Englischen in jeweils eigenen Modellen abgebildet wurde, werden im Folgenden die Zusammenhänge zwischen den beiden sprachlichen Bereichen in einem gemeinsamen Modell dargestellt. Das besondere Interesse betrifft die Korrelationen der latenten Dimensionen auf Klassenebene sowie die Korrelationen zwischen den manifesten Indikatoren der Kompetenzen im Deutschen und im Englischen auf Individualebene (vgl. Abbildung 18.5).

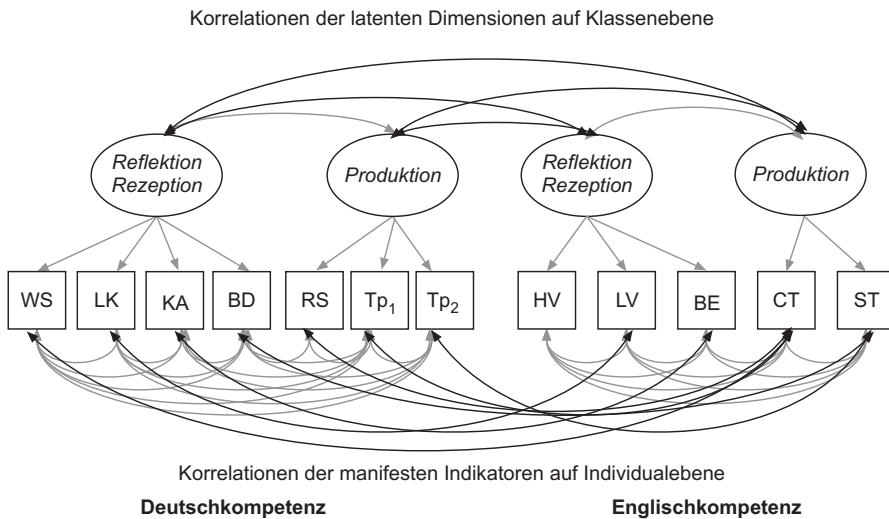


Abbildung 18.5: Gemeinsames Strukturmodell der Deutsch- und Englischkompetenz. Modellpassung: $CFI = 0.99$, $RMSEA = .02$; $SRMR_{within} = .00$, $SRMR_{between} = .02$.

Die Zusammenhänge zwischen den latenten Dimensionen auf Klassenebene sind wie zu erwarten hoch, d.h. Klassen, die gute Leistungen im Deutschen erbringen, weisen diese auch im Englischen auf. Die beiden latenten Dimensionen der Sprachproduktion

im Deutschen und im Englischen korrelieren zu $r = .97$, die reflektiv-rezeptiven Bereiche in beiden Sprachen zu $r = .94$. Auf Individualebene zeigen sich stärker differenzierte Strukturen, die Korrelationen sind in Tabelle 18.4 dargestellt.

Tabelle 18.4: Zusammenhänge der Indikatoren der Englisch- und Deutschkompetenz.

		Englisch Individualebene				
		Hören	Lesen	Bewusst- heit	C-Test	Textpro- duktion
Deutsch Individualebene	Wortschatz	.23	.29	.16	.30	.24
	Lesekompetenz	.24	.31	.18	.32	.27
	Argumentation	.15	.24	.13	.26	.28
	Bewusstheit	.23	.29	.23	.42	.36
	Rechtschreiben	.17	.26	.26	.43	.34
	Textproduktion Pragmatik	.13	.13	.13	.28	.37
	Textproduktion Systematik	.13	.21	.18	.33	.27

Besonders jene Testmodule im Deutschen und im Englischen, die auf ähnliche Bereiche rekurren, weisen auf Individualebene deutliche Zusammenhänge auf. Dazu zählen besonders die Lesekompetenz ($r_{DE} = .31$) und die Module zur Textproduktion ($r_{DE} = .37$) in beiden Sprachen. Hier wurden im Deutschen wie auch im Englischen vergleichbare Aufgaben- und Antwortformate gewählt, so dass Gemeinsamkeiten in den spezifischen Kompetenzen über die Sprachen hinweg auch aus testmethodischer Sicht angenommen werden können (vgl. Beck/Klieme 2007). Hohe Korrelationen mit den Deutschmodulen weist auch der englische C-Test auf. Dieses Verfahren, dass als Indikator der allgemeinen Sprachkompetenz verwendet wird (vgl. Kapitel 14), scheint neben den spezifischen Iteminhalten in der Tat auch mit einer allgemeinen Sprachkompetenz in Zusammenhang zu stehen, die vor allem Sprachbewusstheit und Rechtschreiben, sowie Lesekompetenz und Wortschatz einbezieht.

18.5 Zur Struktur von Sprachkompetenz

Mit den Instrumenten der DESI Studie werden die sprachlichen Kompetenzen wesentlich differenzierter und facettenreicher erfasst, als dies in vergleichbaren Schulleistungsstudien bisher der Fall war. Dies ermöglicht die Analyse der Struktur von Sprachkompetenz, die über eine Modellierung einfacher eindimensionaler Modelle hinausgeht. Durch die Verwendung unterschiedlicher Testformate lassen sich sowohl produktive als auch rezeptive Dimensionen der sprachlichen Kompetenzen im Deutschen wie im Englisch abbilden. Es verwundert daher nicht, dass die vorliegenden Ergebnisse über jene theoretisch postulierten und in Studien der 1970er Jahre auch empirisch gefundenen, eindimensionalen Modelle von Sprachkompetenz hinausgehen. Dies zeigt sich auf der Ebene der individuellen Schülerinnen und Schüler

darin, dass die verschiedenen sprachlichen Kompetenzen zwar miteinander korrelieren, diese Zusammenhänge jedoch durchaus starke individuelle Variationen aufweisen, so dass mehrdimensionale Modelle angemessen abgebildet werden können.

Für das Beispiel der Sprachproduktion bedeutet dies: Schülerinnen und Schüler, deren Rechtschreibkompetenz hoch ausgeprägt ist, müssen nicht zwangsläufig auch gute Aufsätze schreiben. Auch ein großer Wortschatz begünstigt zwar das Leseverstehen, aber alle diese Kompetenzen sind zumindest auf Individualebene aber nicht auf einen gemeinsamen, übergeordneten Faktor zurückzuführen. Die Annahme, dass beispielsweise ein Leseverstehentest als ausschließlicher und ausreichender Indikator für die gesamtsprachliche Kompetenz dienen könnte, ist also nicht haltbar. Vielmehr sollten für Aussagen über sprachliche Leistungen von Schülerinnen und Schülern, die sich auf die individuellen Kompetenzen beziehen, immer facettenreiche Erhebungsinstrumente zum Einsatz kommen. Auch auf Klassenebene zeigt sich ein stärker strukturiertes Kompetenzmodell als in einigen bisherigen Studien.

In DESI lassen sich für das Deutsche und das Englische jeweils zwei latente Dimensionen unterscheiden: Sprachproduktion und Sprachreflektion-Rezeption. Diese Ergebnisse unterstützen die Validität der verwendeten Testinstrumente, die Variationen auch auf Klassenebene noch abbilden lassen. Besonders die Struktur von Sprachkompetenz auf Klassenebene bietet Ansatzpunkte für die Wirkungsforschung des Sprachunterrichts, da die Variationen durchaus auf Unterschiede in den Unterrichtskulturen zurückgeführt werden könnten (De Fraine u.a. 2003). Erst die in DESI realisierte differenzierte Erfassung von Sprachkompetenz als mehrdimensionales Phänomen ermöglicht Rückschlüsse über spezifische Kompetenzen, die auch im Rahmen eines *large scale assessments* durch den Einsatz curriculumsspezifischer, differenzierter Testinstrumente durchaus möglich sind. Zukünftige Studien sollten daher neben der Lesekompetenz, die allgemein als wesentliche Komponente von Sprachkompetenz angesehen wird, auch die Fähigkeiten in weiteren sprachlichen Bereichen im Blick behalten.

Auf Basis dieser differenzierten Testdaten lassen sich in weiteren Analysen zusätzliche Modelle abbilden, die spezifischen theoretischen Ansätzen zur Struktur von Sprachkompetenz nachgehen. Hierzu gehören Zusammenhangsanalysen in Mehrebenenregressionsmodellen und Mehrebenenstrukturmodellen, welche auch die Analyse differenzieller Unterschiede zwischen Personengruppen mit einbeziehen (vgl. Jude im Druck).

Literatur

- Beck, B./Klieme, K. (Hrsg.) (2007): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Walther, G./Valtin, R. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.

- De Fraine, B./Damme, J.v./Landeghem, G.v./Opdenakker, M.-C. (2003): The effect of schools and classes on language achievement. In: *British Educational Research Journal* 29, H. 6, S. 841-859.
- Detering, K (Hrsg.) (2000): *Wortschatz und Wortschatzvermittlung. Linguistische und didaktische Aspekte*. Frankfurt am Main: Lang.
- Deutsches PISA Konsortium (Hrsg.) (2001): *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Eckerth, J. (2000): Zielsprachliche Kommunikation über Grammatik im Fremdsprachenunterricht. In: *Zeitschrift für Fremdsprachenforschung* 11, H. 1, S. 9-30.
- Gräfe-Bentzien, S. (2000): *Evaluierung bilingualer Kompetenz. Eine Pilotstudie zur Entwicklung der deutschen und italienischen Sprachfähigkeiten in der Primarstufe beim Schulversuch der Staatlichen Europa-Schule Berlin (SESB)*. Berlin: Freie Universität.
- Haberzettl, S./Wegener, H. (Hrsg.) (2003): *Spracherwerb und Konzeptualisierung*. Frankfurt am Main: Lang.
- Hox, J.J. (1993): Factor analysis of multilevel data. In: Oud, J.H.L./Blokland-Vogeleesang, R.A.W. (Hrsg.): *Gauging the Muthén Model. Advances in longitudinal and multivariate analysis in the behavioral sciences*. Nijmegen, NL: ITS, S.141-156.
- Hox, J.J. (2002): *Multilevel Analysis. Techniques and Applications*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Jöreskog, K.G. (1993): Testing structural equation models. In: Bollen, K.A./Long, J.S. (Hrsg.): *Testing structural equation models*. Newbury Park, California: Sage Publications, Inc., S. 294-316.
- Jude, N. (im Druck): *Strukturen sprachlicher Kompetenzen im Deutschen und Englischen*.
- Köller, O./Trautwein, U. (2004): Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In: Köller, O./Watermann, R./Trautwein, U./Lüdtke, O. (Hrsg.): *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*, S. 285 – 326.
- Kuhlemeier, H./ van den Bergh, H./Rijlaarsdam, G. (2002): The dimensionality of speaking and writing: A multilevel factor analysis of situational, task and school effects. In: *British Journal of Educational Psychology* 72, S. 467-482.
- Kunnan, A. J. (1992): An investigation of a criterion-referenced test using G-theory, and factor and cluster analyses. In: *Language Testing* 9, H. 1, S. 30-44.
- Kunnan, A. J. (1998): An introduction to structural equation modelling for language assessment research. In: *Language Testing* 15, H. 3, S. 295-332.
- McNamara, T. F. (1995): Modelling performance: opening Pandora's box. In: *Applied linguistics* 16, H. 2, S. 159-175.
- Mehta, P.D./Foorman, B.R./Branum-Martin, L./Taxlor, W.P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. In: *Scientific Studies of Reading* 9, H. 2, S. 85-116.
- Muthén, B. (2004): Latent variable analysis. In: Kaplan, D. (Hrsg.): *Handbook of quantitative methodology for the social science*. Newbury Park, CA, Sage Publications.
- North, B. (2000): *The development of a common framework scale of language proficiency*. Frankfurt am Main: Lang.
- Oerter, R./Montada, L. (2002): *Entwicklungspsychologie*. Weinheim: Psychologie Verlags Union.
- Oller, J. W. (1976): Evidence for a general language proficiency factor: An expectancy grammar. In: *Die Neueren Sprachen* 76, S. 165-174.
- Rost, D. H./ Hartmann, A. (1992): Lesen, Hören, Verstehen. In: *Zeitschrift für Psychologie* 200, S. 345-361.
- Schermelleh-Engel, K./ Moosbrugger, H. (2002). Beurteilung der Modellgüte von Strukturgleichungsmodellen. Frankfurt: Arbeiten aus dem Institut für Psychologie der J. W. Goethe-Universität, Heft 4/2002.

Szagun, G. (2000): Sprachentwicklung beim Kind. Weinheim: Beltz

Tanaka, J. S. (1993): Multifaceted conceptions of fit in structural equation models. In: Bollen, K. A./Long, J. S. (Hrsg): Testing structural equation models. Newbury Park, California: Sage Publications, S. 10-39.

Van den Noortgate, W. M./Opendakker, C./Onghena, O. (2005): The effects of ignoring a level in multilevel analysis. In: School effectiveness and school improvement 16, H. 3, S. 281-303.

Johannes Hartig / Nina Jude

19 Sprachkompetenzen von Mädchen und Jungen

Unterschiedliche Schulleistungen von Jungen und Mädchen in den sprachlichen Fächern werden in vielen aktuellen Leistungsstudien berichtet (z.B. für Lesen in PISA, vgl. Deutsches PISA Konsortium 2001; PISA-Konsortium Deutschland 2004; OECD 2004). Dabei stehen die Unterschiede in Schulleistungen am Ende eines sprachlichen Entwicklungsprozesses, in dem ein Vorsprung der Mädchen schon recht früh zu beobachten ist (Hoff-Ginsberg 2000). Nicht nur die vorschulische Sprachentwicklung – die ersten Sprachversuche, Zuwachs des Wortschatzes – beginnt bei Mädchen früher als bei Jungen, auch Störungen im verbalen Bereich – Stottern, Lese-Rechtschreib-Schwächen – sind bei Schülern wesentlich häufiger zu finden als bei Schülerinnen (Szagun 2000; Bornstein/Hahn/Haynes 2004). Dieser Entwicklungsvorsprung von Mädchen manifestiert sich in der Grundschulzeit besonders im Bereich der Lesekompetenz sowie der Rechtschreibung. Sie zeigen eine schnellere Entwicklung von Rechtschreibstrategien und eine generell höhere Lernmotivation im sprachlichen Kompetenzerwerb (Bos u.a. 2003; Weinert/Helmke 1997). Auf institutioneller Ebene wird dies auch in den Übergangsempfehlungen für die weiterführenden Schulen sichtbar, in denen der Anteil der Mädchen an den höheren Bildungsgängen stetig ansteigt (Baumert/Cortina/Leschinsky 2003). Am Ende der Pflichtschulzeit zeigt sich hinsichtlich der Geschlechtsunterschiede in einigen Studien ein widersprüchliches Bild: Die *Study of Reading Literacy* (Wagemaker 1996) verweist darauf, dass auch noch in der Adoleszenz Entwicklungsvorsprünge aufgeholt werden können. Darüber hinaus erscheint es sinnvoll, spezifische sprachliche Fähigkeiten nach Inhaltsbereichen differenziert zu analysieren (Stanat/Kunter 2001). Diese Möglichkeit bietet sich in DESI, da mit den verwendeten Testinstrumenten insgesamt sieben verschiedene Leistungsbereiche der Unterrichtssprache Deutsch abgedeckt werden. Darüber hinaus erlaubt es DESI, auch Geschlechtsunterschiede in der Fremdsprache Englisch differenziert zu betrachten.

Die Unterschiede zwischen Mädchen und Jungen in den sprachlichen Teilkompetenzen im Deutschen und Englischen am Ende der neunten Jahrgangsstufe werden im Folgenden für die Gesamtstichprobe und differenziert nach Bildungsgängen analysiert. Zusätzlich wird auch auf geschlechtsspezifische Unterschiede im Leistungszuwachs in der neunten Jahrgangsstufe eingegangen.

19.1 Leistungsunterschiede im Deutschen

Im Gesamttestwert für Deutsch finden sich für Mädchen im Mittel 40 Punkte höhere Leistungen als für Jungen. Diese Differenz entspricht einem mittleren Effekt

($d = 0.41$), wie er auch für Lesekompetenz in PISA 2003 berichtet wird ($d = 0.39$; Zimmer/Burba/Rost 2005). In Abbildung 19.1 sind die mittleren Leistungen von Jungen und Mädchen sowie die Effektgrößen der Unterschiede differenziert nach den sieben in DESI erfassten Einzelkompetenzen im Deutschen dargestellt.

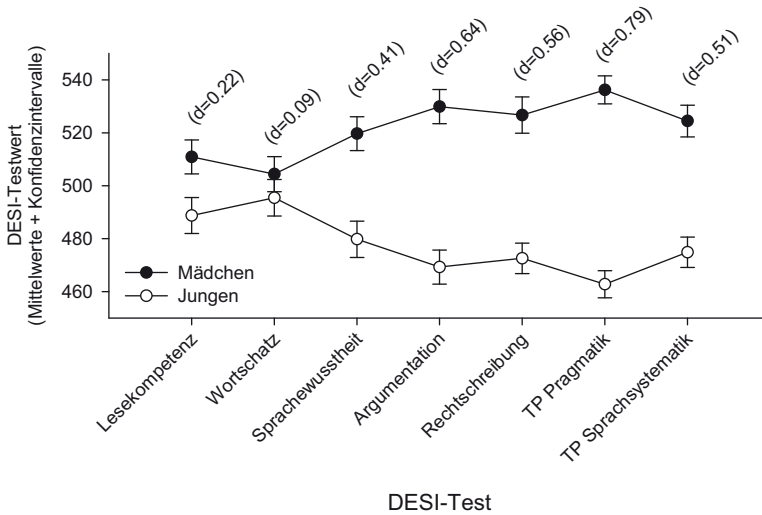


Abbildung 19.1: Leistungen von Mädchen und Jungen in den Teilkompetenzen im Deutschen Ende der neunten Jahrgangsstufe. In Klammern sind die Effektgrößen (d) für die Mittelwertsunterschiede angegeben.

Wie ersichtlich, besteht in allen Kompetenzen im Deutschen ein statistisch bedeutsamer Vorsprung der Mädchen. Während diese Differenz in Wortschatz jedoch sehr gering ist und in Lesekompetenz in der Größenordnung eines kleinen Effektes liegt, liegen die Effektgrößen für die anderen Teilkompetenzen eher im Bereich mittlerer (z.B. für Bewusstheit) bis großer Effekte (für Textproduktion Pragmatik). Die deutlichsten Geschlechtsdifferenzen sind in den produktiven Teilkompetenzen (Textproduktion und Rechtschreibung) sowie in Argumentation zu beobachten. Der Befund für Rechtschreibung deckt sich dabei mit bisherigen empirischen Erkenntnissen anderer Studien, in denen Mädchen in den sprachproduktiven Bereichen auf höheren Kompetenzniveaus durchweg überrepräsentiert sind (vgl. Richter 1994).

19.2 Leistungsunterschiede im Englischen

Auch im Englischen findet sich ein bedeutsamer Gesamtvorsprung der Mädchen; mit 28 Punkten ($d = 0.29$) fällt dieser Unterschied jedoch nicht ganz so stark aus wie im Deutschen. In Abbildung 19.2 sind die Mittelwertsunterschiede zwischen Mädchen und Jungen in den Einzelkompetenzen im Englischen grafisch veranschaulicht.

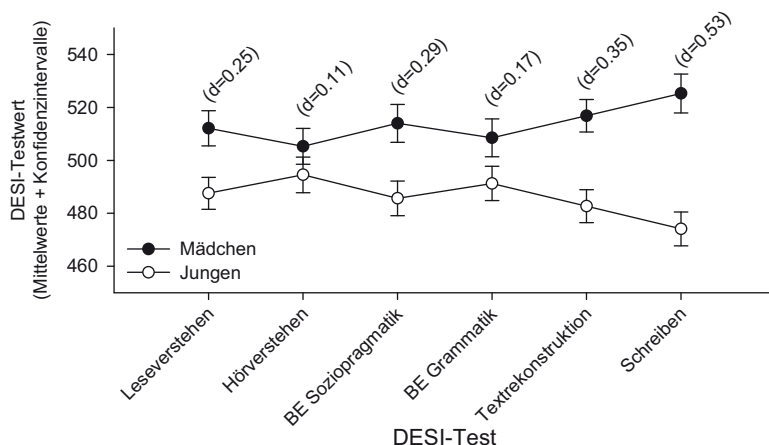


Abbildung 19.2: Leistungen für Mädchen und Jungen in den Teilkompetenzen im Englischen Ende der neunten Jahrgangsstufe. In Klammern sind die Effektgrößen (d) für die Mittelwertsunterschiede angegeben.

Im Englischen ist die Effektgröße der in allen Einzelkompetenzen vorhandenen bedeutsamen Geschlechtsdifferenzen zugunsten der Mädchen insgesamt geringer. Sie reicht von einem praktisch vernachlässigbaren Unterschied im Hörverstehen ($d = 0.11$) bis zu einem mittleren Effekt in Textproduktion ($d = 0.53$). Auch hier sind die größten Vorsprünge für Mädchen in denjenigen Teilkompetenzen zu finden, die sich auf Sprachproduktion beziehen (Textproduktion und C-Test). Diese Ergebnisse decken sich mit bisherigen Befunden, die Geschlechtsunterschiede im Fremdsprachenlernen u.a. auf höhere Motivation und ein größeres Interesse bei Mädchen zurückführen (Coleman 1997; Dresel/Stöger/Ziegler 2005).

19.3 Berücksichtigung der Verteilung auf die Bildungsgänge

Die bereits im Grundschulalter bestehenden Leistungsunterschiede zwischen den Geschlechtern (s.o.) führen zu einer ungleichen Verteilung von Mädchen und Jungen auf die Bildungsgänge im Sekundarbereich. Diese Ungleichverteilung findet sich auch in DESI – der Mädchenanteil beträgt in der Hauptschule 46%, in der Realschule 52%, in der integrierten Gesamtschule 46% und im Gymnasium 55%. Es stellt sich die Frage, inwieweit die in der Gesamtstichprobe beobachteten Leistungsunterschiede zwischen Jungen und Mädchen auf diese Ungleichverteilung zurückzuführen sind und inwieweit die Geschlechtsdifferenzen auch innerhalb der Bildungsgänge bestehen. Eine diesbezügliche Einschätzung erlaubt die nach Bildungsgängen getrennte Betrachtung der Geschlechtsdifferenzen, die in Abbildung 19.3 veranschaulicht sind. Es wird ersichtlich, dass sowohl im Deutschen als auch im Englischen auch innerhalb der Bildungsgänge durchweg signifikante Vorsprünge der Mädchen gegenüber

den Jungen zu beobachten sind, wenngleich diese absolut etwas geringer ausfallen als für die Gesamtstichprobe.

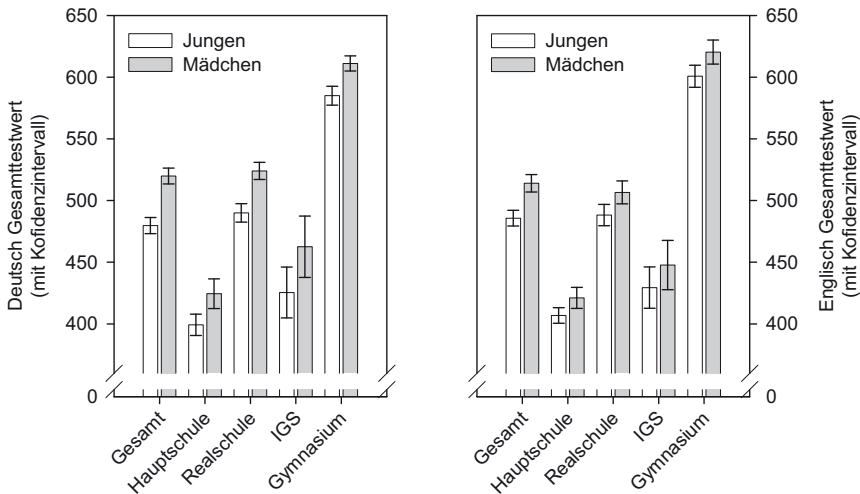


Abbildung 19.3: Mittlere Gesamtestleistungen von Jungen und Mädchen in Deutsch und Englisch in den verschiedenen Bildungsgängen.

19.4 Leistungszuwächse von Jungen und Mädchen

Auch hinsichtlich des Leistungszuwachses in der neunten Jahrgangsstufe zeigt sich für die DESI-Stichprobe ein Bild, das Ergebnisse anderer Studien widerspiegelt. Mädchen zeigen in allen Sprachkompetenzen (außer in der Textproduktion Systematik) einen signifikanten Leistungszuwachs, während dieser bei Jungen nur in der Sprachbewusstheit Deutsch sowie dem Hörverstehen und dem C-Test im Englischen festgestellt werden kann (vgl. Abbildung 19.4). Hinzu kommt, dass bei Schülerinnen ein stärkerer Zuwachs in allen sprachlichen Bereichen zu verzeichnen ist als bei den Schülern. Auch dieser Befund steht in Einklang mit Ergebnissen anderer Längsschnittstudien der Lernentwicklung in sprachlichen Bereichen (Lehmann u.a. 2004).

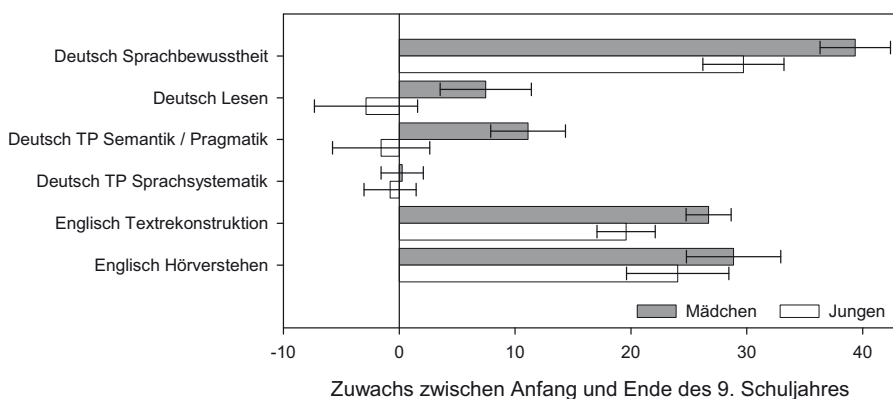


Abbildung 19.4: Leistungsdifferenzen zwischen Anfang und Ende der neunten Jahrgangsstufe für Mädchen und Jungen im Deutschen und Englischen.

19.5 Geschlechtsdifferenzen in Schulklassen

Über die hier berichteten Mittelwertsunterschiede hinaus stellt sich die Frage, ob sich Schulklassen hinsichtlich der Geschlechtsdifferenzen unterscheiden, oder ob die beobachteten Differenzen über unterschiedliche Klassen konsistent sind. Dieser Frage wurde in Mehrebenenanalysen nachgegangen. Es zeigt sich, dass die Unterschiede zwischen Jungen und Mädchen auch unter Kontrolle des Bildungsganges signifikant variieren ($\chi^2 = 538.9$; $df = 403$; $p < .001$ für Gesamtleistung Deutsch und $\chi^2 = 607.0$; $df = 403$; $p < .001$ für Gesamtleistung Englisch). Es erscheint für differenzierte Analysen also durchaus lohnend, nach erklärenden Variablen auf Klassen- oder Unterrichtsebene zu suchen, die das Zustandekommen von Geschlechtsdifferenzen fördern oder vermindern. Hierzu zählen Überzeugungen und Wertesysteme von Lehrpersonen ebenso wie geschlechtsspezifische Interaktionsstile (Dresel/Stöger/Ziegler 2005). Es kann festgehalten werden, dass das Zustandekommen von Geschlechtsunterschieden in den sprachlichen Kompetenzen hinsichtlich vielfältiger, möglicher Bedingungsfaktoren weiter untersucht werden sollte.

Literatur

- Baumert, J./Cortina, K.S./Leschinsky, S. (2003): Grundlegende Entwicklungen und Strukturprobleme im allgemein bildenden Schulwesen. In: Cortina, K.S./ J. Baumert, J./Leschinsky, A./Mayer, K.U./Trommer, L (Hrsg.): Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick. Reinbek: Rowohlt, S. 52-147
- Bornstein, M.H./ Hahn, C.-S./Haynes, O.M. (2004): Specific and general language performance across early childhood: Stability and gender considerations. In: *First language* 24, H. 3, S. 267-304.

- Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Walther, G./Valtin, R. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Waxmann: Münster.
- Cohen, J. (²1988): Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum.
- Coleman, J. (1997): Residence abroad within language study. In: Language Teaching 30, H.1, S. 1-20.
- Deutsches PISA Konsortium (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Dresel, M./Stöger, H./Ziegler, A. (2005): Klassen- und Schuleffekte bei Geschlechtsunterschieden im schulischen Leistungsbereich (Ulmer Forschungsberichte aus der Pädagogischen Psychologie Nr. 12). Ulm: Universität, Abt. Pädagogische Psychologie.
- Hoff-Ginsberg, E. (2000): Soziale Umwelt und Sprachlernen. In: Grimm, H (Hrsg.): Sprachentwicklung. Enzyklopädie der Psychologie, C, III, 3. Göttingen: Hogrefe, S.463-494.
- Lehmann, R./Hunger, S./Ivanov, S./Gänsfuß, R./Hoffmann, E. (2004): Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 11. Ergebnisse einer Längsschnittuntersuchung in Hamburg. Hamburg: Behörde für Bildung und Sport
- OECD (2004): Lernen für die Welt von morgen. Erste Ergebnisse von PISA 2003. Paris: OECD.
- PISA-Konsortium Deutschland (Hrsg.) (2004): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann.
- Richter, S. (1994): Geschlechtsunterschiede in der Rechtschreibentwicklung von Kindern der 1. bis 5. Klasse. In Richter, S./Brügelmann, H. (Hrsg.): Mädchen lernen anders – anders lernen Jungen: Geschlechtsspezifische Unterschiede beim Schriftspracherwerb. Bottighofen: Libelle, S. 51-65.
- Stanat, P./Kunter, M. (2001): Geschlechtsunterschiede in Basiskompetenzen. In: Deutsches PISA Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich: Opladen: Leske + Budrich, S.251-270.
- Szagun, G. (2000): Sprachentwicklung beim Kind. Weinheim: Beltz
- Wagemaker, H. (Hrsg.) (1996): Are girls better readers? Gender differences in reading literacy in 32 countries. Den Haag: IEA.
- Weinert, F.E./Helmke, A. (1997): Entwicklung im Grundschulalter. Weinheim: Beltz.
- Zimmer, K./Burba, D./Rost, J. (2005): Kompetenzen von Jungen und Mädchen. In: PISA-Konsortium Deutschland (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs .Münster: Waxmann, S.211-223.

Hermann-Günter Hesse / Kerstin Göbel / Johannes Hartig

20 Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache¹

Die europäischen politischen Institutionen haben eine Reihe von Maßnahmen in Gang gesetzt mit dem Ziel, dass jeder europäische Bürger in naher Zukunft drei Sprachen beherrschen soll – neben der Erstsprache zwei weitere. Mehrsprachigkeit sei eine Schlüsselqualifikation sowohl für den Einzelnen als auch für die Gesellschaft (Kommission der Europäischen Gemeinschaften 2005). 19% der Schülerinnen und Schüler der DESI-Studie haben ihren Angaben zufolge eine andere als die deutsche Sprache als Erstsprache erworben oder bezeichnen eine weitere Sprache oder mehrere neben Deutsch als ihre Erstsprache. Diese Gruppe lebt unter Bedingungen, die ihnen das Erlernen mehrerer Sprachen nahelegt oder – je nach Sichtweise – abverlangt. Wenn sie in diesen Sprachen auf einem intermediären Niveau – so das Kriterium der Europäischen Kommission – kompetent sein sollten, dann hätten sie das europäische Ideal bereits erreicht.

Bisher standen in groß angelegten Studien wie TIMSS oder PISA kaum sprachliche Leistungen im Vordergrund; es wurde lediglich die Lesekompetenz in der Unterrichtssprache untersucht. Die diesbezüglichen Ergebnisse zeigen für Deutschland regelmäßig einen starken Zusammenhang von Migrationsstatus, sozialer Herkunft und Schülerleistungen (z.B. OECD 2006). Lesekompetenz ist aber als Indikator für eine Einschätzung der allgemeinen Sprachfähigkeit nicht ausreichend (vgl. Kapitel 18). Über den Erwerb der schulisch ersten Fremdsprache liegen bisher keine Erkenntnisse aus Large-Scale-Assessments vor. Somit stellt sich die Frage, ob sich der beobachtete Trend fortsetzt, wenn sprachliche Leistungen über Lesekompetenz hinaus thematisiert werden und wie es sich beim Erlernen der schulisch ersten Fremdsprache verhält – für viele Einwandererkinder bereits der dritte oder vierte Spracherwerb.

In den folgenden Abschnitten soll auf die folgenden Fragen eingegangen werden: Stellen mehrsprachige Lernumwelten für die Leistungen in den Fächern Deutsch und Englisch günstige oder ungünstige Bedingungen dar (20.1)? Gibt es einen Zusammenhang zwischen der Sprachvitalität in der Erstsprache und dem Erwerb weiterer Sprachen (20.2)? Wie wirkt sich die Zusammensetzung der Schülerschaft in einer Schulklasse in Bezug auf mono- und multilinguale Schülerinnen und Schüler auf sprachliche Leistungen aus (20.3)?

1 Für die hilfreiche Kommentierung und Kritik bedanken wir uns bei Dominique Rauch, DIPF.

20.1 Lernumwelten und sprachliche Kompetenzen

Vor dem Hintergrund der verschiedenartigen Ergebnisse zur Auswirkung des Mehrsprachenerwerbs auf die Beherrschung einzelner Sprachen fällt es schwer, Hypothesen über zu erwartende Ergebnisse abzuleiten. Den Thesen der nachteiligen Wirkungen, insbesondere des frühen Zweit- oder Mehrsprachenerwerbs (Weisgerber 1966) stehen Postulate positiver Effekte gegenüber (Cummins 1993; Gegenüberstellungen bei Kielhöfer/Jonekeit 1983). Auch wenn man rückblickend die Entwicklung der Einstellungen gegenüber den Wirkungen der Mehrsprachigkeit als die Abfolge von behaupteten negativen, positiven und neutralen Effekten beschreiben kann (Bialystok 1991), so sind die jeweils verfochtenen widersprüchlichen Thesen bisher kaum mittels wissenschaftlich fundierter, standardisierter Testverfahren in größer angelegten oder gar repräsentativen Studien geprüft worden.

Die frühen Forschungen zur Mehrsprachigkeit waren von der Konfundierung von Bedingungen der sprachlichen Lernumwelt und der sozialen Herkunft der untersuchten Personen geprägt: Mehrsprachige mit einem eher niedrigen sozioökonomischen Hintergrund wurden mit Einsprachigen aus den mittleren und höheren Einkommensgruppen verglichen, ohne diesen Umstand bei der Analyse zu berücksichtigen. Wurden sozioökonomische Variablen nicht ausreichend in Rechnung gestellt, bestand die Gefahr, falsche Schlüsse zu ziehen, denn die Zugehörigkeit zu einer unteren sozioökonomischen Schicht führte offenbar zum Ergebnis negativer Wirkungen der Mehrsprachigkeit. Seither ist die Berücksichtigung der familiären Hintergrundvariablen eine wesentliche Forderung an die Interpretation von Forschungsergebnissen zum Zweit- und Mehrsprachenerwerb.

Interkulturell orientierte Untersuchungen, welche die Konfundierung der sprachlichen Leistungen mit den familiären Charakteristiken kontrollieren oder die so gestaltet wurden, dass sie keine Rolle spielt, teilen der Mehrsprachigkeit die Bedeutung einer die metasprachlichen Fähigkeiten fördernde Bedingung zu (Mohanty/Perregaux 1997; Segalowitz 1981). Mehrsprachige erfahren demnach bereits früh die semantische Relativität, dies fördere die Beherrschung kognitiver Kontrollprozesse verbunden mit erhöhter Aufmerksamkeit. Diese kognitiven Vorteile der Mehrsprachigkeit sollten den weiteren Fremdsprachenerwerb erleichtern (Bialystok 2005). Was jedoch die empirischen Belege angeht, gibt es bislang keineswegs eine einheitliche Antwort auf die Frage, ob Bilinguale eine dritte Sprache schneller oder besser lernen als Monolinguale eine zweite Sprache (Cenoz 2003). Wenn die These der spracherwerbsbezogenen Förderung kognitiver Fähigkeiten zutrifft, dann sollten simultan und sukzessiv Mehrsprachige beim Erwerb einer zusätzlichen Sprache erfolgreicher sein als monolingual Aufgewachsene – vorausgesetzt die genannten konfundierenden Variablen werden kontrolliert.

Die Schulleistungstudien der letzten Jahre haben so gut wie durchgängig die problematische Lage von Migranten insbesondere in Deutschland gezeigt (OECD 2003). Der UNO-Inspektor des Menschenrechtsrats, Vernor Muñoz, wirft Deutschland vor, Migranten zu benachteiligen, sein Bildungssystem verwehre ih-

nen die Chancengleichheit (United Nations 2007). Die dauerhafte Koppelung des Migrationshintergrunds mit einem niedrigen sozialen Status wird im Bildungsbericht für Deutschland als institutionelle Segregation der eingewanderten Bevölkerung und damit als eine besondere Herausforderung für die Bildungspolitik bezeichnet (Konsortium Bildungsberichterstattung 2006). Wenn auch die Abhängigkeit der Schülerleistungen vom Migrationshintergrund und der sozialen Herkunft in vielen – nicht allen – Ländern gegeben ist, so zeigen doch die internationalen Vergleiche, dass diese Abhängigkeit in Deutschland besonders stark ausgeprägt ist (OECD 2006).

Erwartet man einerseits erhöhte Spracherwerbsleistungen simultan und sukzessiv Mehrsprachiger und geht man andererseits davon aus, dass diese Schülerinnen und Schüler in Lernumwelten heranwachsen, die für das Zustandekommen guter schulischer Leistungen ungünstige Voraussetzungen bereitstellen, dann ist es unabdingbar, das Zusammenwirken von Sprachbiographie und Lernumwelt zum Gegenstand der Analyse zu machen.

Erstsprachen und Migrationshintergrund

Die Bedeutung, die in neuester Zeit dem Drittsprachenerwerb von simultan Bilingualen zugesprochen wird (Cenoz/Hufeisen/Jessner 2001; Sanz 2000), legt es nahe, eine differenziertere Kategorisierung der Personen auf Grund ihrer Sprachbiographie vorzunehmen. Es gibt Hinweise darauf, dass sich der Drittsprachenerwerb von simultan Bilingualen von dem Monolingualer oder sukzessiv Bilingualer unterscheidet. Deshalb ist es zweckmäßig, die Gruppe der Schülerinnen und Schüler, die (auch) eine andere Sprache als Deutsch als Erstsprache erworben haben, entsprechend zu differenzieren. Diese Differenzierung ist in bisherigen Large-Scale-Studien nicht vorgenommen worden. Im DESI-Projekt wurden die Schülerinnen und Schüler gefragt, ob sie neben Deutsch noch eine oder mehrere andere Sprachen als Erstsprachen erworben haben, wenn ja, welche. Bei positiver Beantwortung dieser Frage wird davon ausgegangen, dass damit ein simultaner Bi- oder Multilingualismus (L1/L2 oder L1/L2/L3) vorliegt, wobei eine der Sprachen Deutsch ist. Sie werden im Folgenden als „Mehrsprachige“² bezeichnet. Sie stellen 6% der DESI-Gesamtstichprobe dar (vgl. Tabelle 20.1). 13% geben ausschließlich eine andere Erstsprache als Deutsch an. Von diesen Schülerinnen und Schülern, die zunächst monolingual mit einer anderen Sprache als Deutsch aufgewachsen sind, ist anzunehmen, dass sie spätestens mit dem Eintritt in die deutsche Schule, unter Umständen aber auch schon im Kindergarten, Deutsch als zweite Sprache erworben haben; sie zeichnen sich somit durch einen sukzessiven Spracherwerb aus (L1 → L2), wobei die L1 unterschiedlich, je nach Herkunftsfamilie, die L2 Deutsch ist. Diese sukzessiv Bilingualen werden im Folgenden zur besseren Abgrenzung als „nicht deutsch Erstsprachige“ bezeichnet. Damit wäre etwa ein Drittel derer, die nicht oder nicht ausschließlich Deutsch als

2 Es sei darauf hingewiesen, dass der Begriff „Mehrsprachigkeit“ in der Literatur nicht einheitlich gebraucht wird. Unter „Mehrsprachigen“ werden auch Personen bezeichnet, die irgendwann in ihrem Lebenslauf eine weitere Sprache erworben haben.

Erstsprache nennen, simultan mehrsprachig aufgewachsen. Die größte Gruppe stellt 81% der Schülerinnen und Schüler dar, die ausschließlich Deutsch als Erstsprache angeben.

Tabelle 20.1: Relative Anzahl der Schülerinnen und Schüler unterschiedlicher Erstsprache in Bezug auf den Migrationshintergrund in Prozent.

Erstsprache	Migrationshintergrund				
	Erste Generation	Zweite Generation	Ein Elternteil nicht in D geboren	Beide Eltern in D geboren	
Deutsch	1	1	6	92	100
Nicht deutsch	54	32	8	5	100
Mehrsprachig	19	28	32	21	100

Erwartungsgemäß liegt auch in der DESI-Studie ein deutlicher Zusammenhang von Mehrsprachigkeit und Migrationsstatus vor. 90% der Schülerinnen und Schüler, die angeben, eine andere Sprache als Deutsch oder neben Deutsch als Erstsprache erworben zu haben, hat einen Migrationshintergrund – im Vergleich zu 8% der Schüler deutscher Erstsprache. Für die Beschreibung des Einwanderungsstatus wurde analog zu PISA (OECD 2003) ein vierstufiger Indikator mit den Ausprägungen „beide Eltern in Deutschland geboren“ (entspricht „kein Migrationshintergrund“), „ein Elternteil nicht in Deutschland geboren“, „Zweite Generation“ und „Erste Generation“ gebildet. Die erste Generation umfasst Schüler, die selbst wie ihre Eltern im Ausland geboren wurden; als zweite Generation werden die Schüler bezeichnet, die selbst in Deutschland, deren Eltern aber im Ausland geboren wurden. Tabelle 20.1 zeigt den Zusammenhang zwischen der Erstsprache der Schülerinnen und Schüler und ihrem Migrationshintergrund. In der Gruppe der Deutschsprachigen finden sich kaum Schülerinnen und Schüler mit Migrationshintergrund (8%), während umgekehrt der weitaus größte Teil der Schülerinnen und Schüler (95%) mit ausschließlich einer nicht deutschen Erstsprache einen Migrationshintergrund aufweist. Die Mehrsprachigen verteilen sich weit gleichmäßiger über die Kategorien des Migrationshintergrunds.

Die Schülerinnen und Schüler nicht deutscher Erstsprache und die Mehrsprachigen unterscheiden sich in Bezug auf die erworbenen Sprachen (vgl. Tabelle 20.2). Türkisch und Russisch berichtet mehr als die Hälfte derer, die eine andere Erstsprache als Deutsch erworben haben, im Vergleich zu nur etwa einem Drittel bei den Mehrsprachigen. Türkisch überwiegt auch bei den Mehrsprachigen neben Polnisch, Südslawisch und insbesondere Englisch. Etwa viermal so viele Mehrsprachige haben im Vergleich zu den nicht deutsch Erstsprachigen bereits zu Hause Englisch gelernt.

Tabelle 20.2: Relative Anzahl der Schüler unterschiedlicher Erstsprache in Bezug auf die Sprachgruppen in Prozent.

Erstsprache	Sprachgruppen						
	Türkisch	Russisch	Polnisch	Süd-slawisch	Englisch	Andere	
Nicht deutsch	27	27	8	3	4	31	100
Mehrsprachig	17	13	11	6	15	38	100

Erstsprache und Bildungsgang, sozioökonomischer Status und allgemeine kognitive Grundfähigkeiten

Schülerinnen und Schüler nicht deutscher Erstsprache besuchen häufiger eine Hauptschule (47%) oder Integrierte Gesamtschule (11%) als die beiden anderen Gruppen und seltener eine Realschule oder ein Gymnasium (vgl. Tabelle 20.3). Die größten Anteile an Gymnasiasten und Realschülern finden sich in der Gruppe der deutschsprachigen Schüler, von denen nur verhältnismäßig wenige eine Hauptschule (23%) oder eine IGS (9%) besuchen. Der relative Anteil der mehrsprachigen Schüler ist im Gymnasium fast so hoch wie jener der deutschsprachigen Schüler und etwa doppelt so hoch wie der der Schüler nicht deutscher Erstsprache. Bei den anderen Bildungsgängen liegt der Anteil der mehrsprachigen Schüler jeweils zwischen dem der anderen beiden Gruppen (Hauptschule: 30%, IGS: 9%, Realschule: 32%).

In Bezug auf die Bildungsgänge sind die Deutscherstsprachigen überproportional in den Gymnasien, die ausschließlich nicht deutsch Erstsprachigen hingegen häufig in der Hauptschule anzutreffen. Die Mehrsprachigen verteilen sich näherungsweise proportional auf die Bildungsgänge mit einem leichten Überhang in der Hauptschule.

Tabelle 20.3: Relative Anzahl der Schüler unterschiedlicher Erstsprache in Bezug auf den Bildungsgang in Prozent.

Erstsprache	Bildungsgang				
	Hauptschule	IGS	Realschule	Gymnasium	
Deutsch	22	9	38	31	100
Nicht deutsch	47	11	27	15	100
Mehrsprachig	30	9	32	29	100

Die Tabelle 20.3 weist auf die ungleiche Verteilung der Schülerinnen und Schüler verschiedener Erstsprache auf die Bildungsgänge hin. Für die Interpretation der Ergebnisse bedeutet dies, dass die Effekte der potenziell für die Sprachleistungen verantwortlichen Faktoren in bivariater Betrachtungsweise kaum abzuschätzen sind. Es bedarf modellorientierter Analysen, um solche Konfundierungen aufzulösen (Baumert/Stanat/Watermann 2006). Dennoch sollen in den beiden folgenden Abschnitten zunächst die Leistungsergebnisse im Sinne einer reinen Bestandsaufnahme, lediglich unter Berücksichtigung der Bildungsgänge dargestellt werden.

Deutsch Gesamtestleistungen der Schüler verschiedener Erstsprachen

In der Abbildung 20.1 sind die Mittelwerte der Deutsch Gesamtleistung (mit Konfidenzintervallen, vgl. Kapitel 3) am Ende der neunten Jahrgangsstufe für die drei auf Grund der Erstsprachen eingeteilten Schülergruppen dargestellt; zunächst über alle Bildungsgänge hinweg (am linken Rand der Abbildung) und nach Bildungsgängen aufgegliedert. Über alle Bildungsgänge hinweg erreichten die Schülerinnen und Schüler, deren Erstsprache ausschließlich Deutsch ist, den höchsten mittleren Gesamtwert für die Deutsch-Leistungstests. Das Testergebnis der Mehrsprachigen ist signifikant geringer als das der deutsch Erstsprachigen, aber signifikant höher als das der Schüler nicht deutscher Erstsprache. Wird der Bildungsgang berücksichtigt, bleibt die Abfolge der Testmittelwerte der Gruppen zueinander erhalten. Die Mittelwertsunterschiede zwischen den deutsch Erstsprachigen und den mehrsprachigen Schülerinnen und Schüler sind nur bei denen, die eine Hauptschule oder ein Gymnasium besuchen, signifikant. In der Realschule und in der Integrierten Gesamtschule unterscheiden sich die Deutsch-Gesamtleistungen der deutsch Erstsprachigen und der Mehrsprachigen nicht signifikant voneinander. Die deutsch Erstsprachigen und die Mehrsprachigen schneiden auch unter Berücksichtigung des Bildungsganges jeweils signifikant besser ab, als die Schüler nicht deutscher Erstsprache. Der Mittelwertsunterschied zwischen Schülern deutscher und nicht deutscher Erstsprache beträgt in der Hauptschule 55, in der Realschule 44, in der Integrierten Gesamtschule 85 und im Gymnasium 57 Punkte. Dies sind erhebliche Unterschiede, die den deutlichen Rückstand der nicht deutsch Erstsprachigen im Fach Deutsch verdeutlichen.

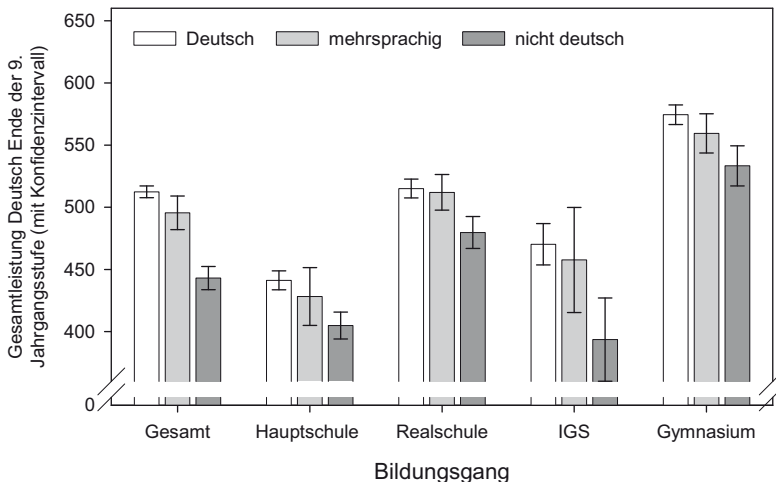


Abbildung 20.1: Erstsprache, Bildungsgang und Deutschgesamtestleistungen.³

3 Die in der Abbildung eingezeichneten Konfidenzintervalle stellen Intervalle um die Mittelwerte und nicht um die Differenz der Mittelwerte dar. Mittelwertunterschiede können signifikant sein obwohl sich Konfidenzintervalle überschneiden.

Betrachtet man den Leistungszuwachs im Deutschen, so ergibt sich nur in der Teilttestleistung ‚Bewusstheit Deutsch‘ im Gymnasium ein statistisch bedeutsamer Anstieg. Die nicht deutsch Erstsprachigen erzielen am Gymnasium im Kompetenzbereich Bewusstheit Deutsch einen Zuwachs von 45 Punkten, die deutsch Erstsprachigen von 43 Punkten. In beiden Fällen ist der Zuwachs signifikant größer als bei den Mehrsprachigen, die nur 27 Punkte hinzu gewinnen.

Englisch Gesamtttestleistungen der Schüler verschiedener Erstsprachen

Die zur vorherigen Abbildung 20.1 analog aufgebaute Abbildung 20.2 zeigt die Testmittelwerte für die Englisch Gesamtleistungen. Im Unterschied zur Deutschleistung sind es im Englischen die Mehrsprachigen, die zum Ende des neunten Schuljahres in allen Bildungsgängen Leistungsmittelwerte erzielen, die höher sind als die der beiden anderen Gruppen. Ihr Abstand zu den deutsch Erstsprachigen ist jedoch nur in der Realschule mit 17 Punkten signifikant. Nimmt man alle Bildungsgänge zusammen, dann sind die Unterschiede der Englischleistungen zwischen den Mehrsprachigen bzw. den deutsch Erstsprachigen und den Leistungen der nicht deutsch Erstsprachigen signifikant. Während die Leistungsdifferenzen zwischen den drei Schülergruppen in der Hauptschule sehr gering ausfallen, zeigt sich in der Integrierten Gesamtschule ein deutlicher Leistungsabstand zwischen den nicht deutsch Erstsprachigen und den beiden anderen Gruppen. Im Gymnasium ist einzig der Abstand zwischen den Leistungen der Mehrsprachigen und den Leistungen der nicht deutsch Erstsprachigen in Höhe von 19 Punkten signifikant.

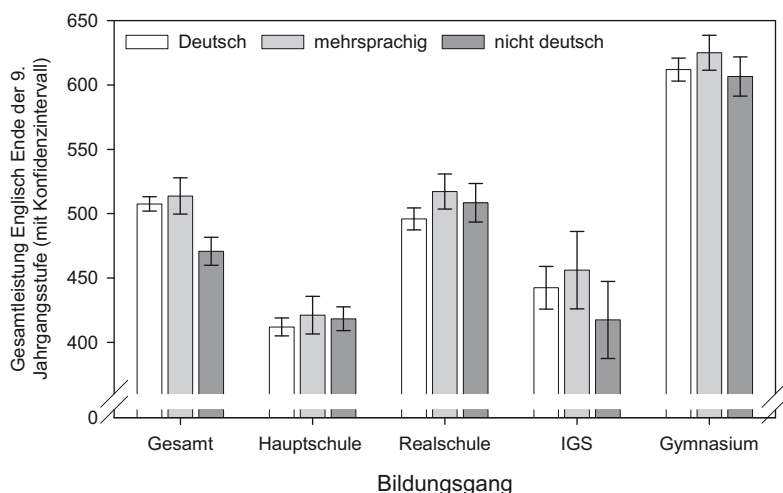


Abbildung 20.2: Erstsprache, Bildungsgang und Englischgesamtttestleistungen.⁴

In den in Abbildung 20.2 dargestellten Mittelwerten sind die Leistungen derjenigen Schülerinnen und Schüler enthalten, die angaben, Englisch als Erstsprache erwor-

⁴ Siehe vorherige Fußnote.

ben zu haben. Um zu prüfen, ob dadurch die Ergebnisse verfälscht werden würden, sind die Gesamtleistungen im Englischen ohne die englisch Erstsprachigen über die Bildungsgänge hinweg berechnet worden. Tatsächlich sinken die Mittelwerte der nicht deutschsprachigen Gruppe und der Gruppe der Mehrsprachigen, jedoch nur um jeweils 5 Punkte. Die Mehrsprachigen erreichen dann im Mittel drei Punkte und die nicht Deutschsprachigen 45 Punkte weniger als die deutsch Erstsprachigen. Das Muster der Mittelwertsunterschiede und die Ergebnisse ihrer inferenzstatistischen Absicherung bleiben hiervon unberührt: Nach dem Ausschluss der englisch Erstsprachigen bleiben nach wie vor die Mittelwertsunterschiede in der Englischgesamtleistung zwischen den deutsch Erstsprachigen und den Mehrsprachigen relativ gering und sind nicht signifikant, die Mittelwertsunterschiede zwischen den nicht deutsch Erstsprachigen und den anderen Gruppen sind dagegen deutlich größer und signifikant.

Im Rahmen der Analyse des differentiellen Lernzuwachses vom Beginn bis zum Ende des neunten Jahrgangs über alle Bildungsgänge hinweg ergibt sich beim C-Test ein signifikanter Unterschied. Die deutsch Erstsprachigen erreichen einen Zuwachs von immerhin 24 Punkten. Demgegenüber ist bei den nicht Deutschsprachigen nur einen Zuwachs von 19 Punkten festzustellen. Im C-Test haben die Mehrsprachigen mit nur 17 Punkten auch einen signifikant geringeren Zuwachs als die deutsch Erstsprachigen. Weiterhin findet sich in Deutsch Lesekompetenz ein größerer Zuwachs für die deutsch Erstsprachigen als für die nicht deutsch Erstsprachigen. Darüber hinaus liegen keine signifikanten differentiellen Zuwächse vor.

Modellorientierte Betrachtung der differentiellen Testleistungen im Deutschen und Englischen

Die bislang dargestellten Ergebnisse veranschaulichen die Verteilungen der Testleistungen, so wie sie unmittelbar in den einzelnen Bildungsgängen und ohne Berücksichtigung von nicht oder kaum beeinflussbaren sozialen und individuellen Merkmalen anzutreffen sind. Schon in der rein deskriptiven Darstellung der Testergebnisse fällt der relativ hohe Leistungsstand der Mehrsprachigen auf.

Interessiert jedoch die Frage, inwieweit es tatsächlich einen Zusammenhang zwischen dem sprachlichem Hintergrund und den Leistungen im Deutschen und Englischen gibt, dann sollten all diejenigen Faktoren statistisch kontrolliert werden, die mit der Erstsprache konfundiert sind. Im Folgenden werden daher die relevanten Merkmale in einer Regressionsanalyse als Kontrollvariablen einbezogen (vgl. Kap. 3). In Abbildung 20.3 sind entsprechend dieser Analysen die Effekte des sprachlichen Hintergrundes auf die Kompetenzen im Deutschen und Englischen dargestellt, die sich bei gleichzeitiger Kontrolle von Bildungsgang, sozioökonomischem Hintergrund, kognitiver Grundfähigkeit und Geschlecht ergeben.

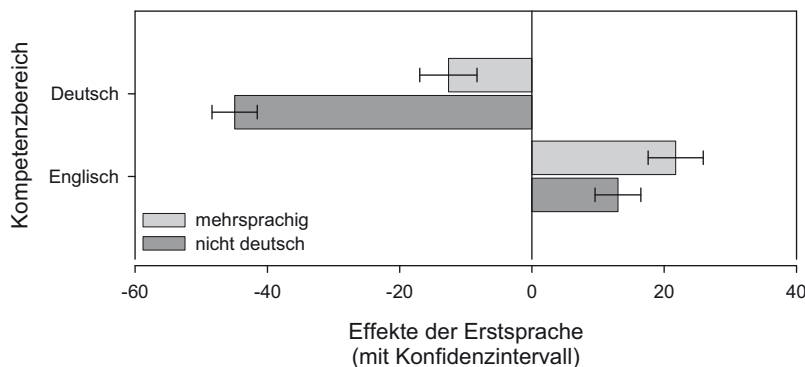


Abbildung 20.3: Effekte des sprachlichen Hintergrundes auf Kompetenzen im Deutschen und Englischen zum Ende der neunten Jahrgangsstufe unter Kontrolle von Bildungsgang, sozioökonomischem Hintergrund, kognitiver Grundfähigkeit und Geschlecht: Kontraste der Leistungen von Mehrsprachigen und nicht deutsch Erstsprachigen gegenüber deutsch Erstsprachigen.⁵

Die Grafik in Abbildung 20.3 bezieht die Testleistungen der nicht deutsch Erstsprachigen und der Mehrsprachigen auf die der deutsch Erstsprachigen. Die mittlere Testleistung der deutsch Erstsprachigen ist entsprechend als Null-Linie eingezeichnet. Noch deutlicher als in den rein deskriptiven Darstellungen der Abbildungen 20.1 und 20.2 zeigen sich die signifikant niedrigen Leistungen der Mehrsprachigen und der nicht deutsch Erstsprachigen im Kompetenzbereich Deutsch und die im Vergleich zu den deutsch Erstsprachigen signifikant höheren Leistungen dieser Gruppen im Kompetenzbereich Englisch. Besonders drastisch fällt die niedrige mittlere Deutsch-Testleistung der nicht deutsch Erstsprachigen mit 45 Punkten gegenüber den deutsch Erstsprachigen aus. Bei den Mehrsprachigen ist der Abstand zu den deutsch Erstsprachigen zwar auch signifikant, aber deutlich geringer.

In Bezug auf Englisch zeigt sich nun, bei Kontrolle der konfundierten Variablen, noch deutlicher als in den deskriptiven Analysen die Überlegenheit der Testleistungen der Mehrsprachigen gegenüber denen der deutsch Erstsprachigen. Aber auch die Englischleistungen der nicht deutsch Erstsprachigen liegen signifikant über denen der deutsch Erstsprachigen.

Daraus kann folgender Schluss gezogen werden: Sieht man von den sich überlagernden Auswirkungen der Bildungsgänge, des sozioökonomischen Hintergrunds, der kognitiven Grundfähigkeiten und des Geschlechts auf die sprachlichen Leistungen ab, dann sind die Testleistungen derjenigen Schülerinnen und Schüler, die ausschließlich oder zusätzlich eine andere Erstsprache als Deutsch erworben haben im Englischen

5 Anders als in den Abbildungen 20.1 und 20.2 beziehen sich die Konfidenzintervalle in den Abbildungen 20.3 bis 20.5 auf Regressionskoeffizienten. In diesen Grafiken ist es möglich, signifikante Effekte daran zu erkennen, dass die dargestellten Konfidenzintervalle die Null-Linie nicht einschließen.

deutlich besser, im Deutschen hingegen deutlich schlechter als die der Schülerinnen und Schüler, die nur Deutsch als Erstsprache erworben haben.

Wiederholt sich dieses globale Ergebnis bei der Analyse der einzelnen durch DESI erfassten Kompetenzbereiche? Um diese Frage zu beantworten, wurden für alle Kompetenzbereiche im Deutschen und im Englischen die zuletzt dargestellten Analysen analog durchgeführt. Die Ergebnisse sind in den Abbildungen 20.4 und 20.5 dargestellt.

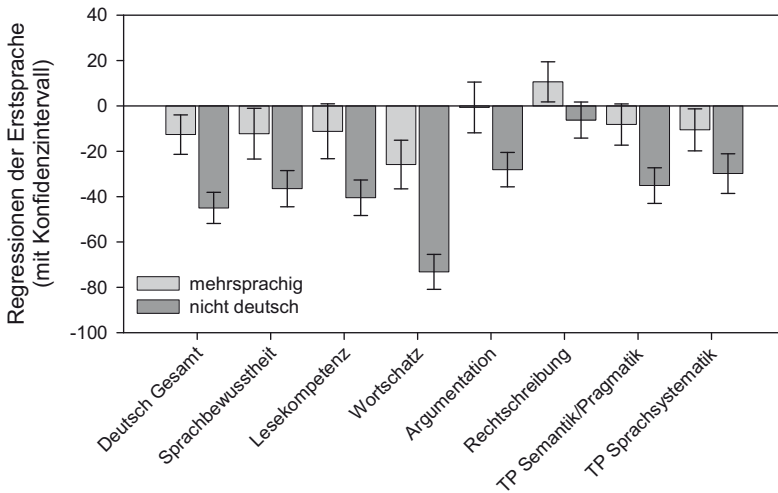


Abbildung 20.4: Effekte des sprachlichen Hintergrundes auf Kompetenzbereiche im Deutschen zum Ende des neunten Jahrgangs unter Kontrolle von Bildungsgang, sozioökonomischem Hintergrund, kognitiver Grundfähigkeit und Geschlecht: Kontraste der Leistungen von Mehrsprachigen und nicht deutsch Erstsprachigen gegenüber deutsch Erstsprachigen.

Die Analysen für die einzelnen Kompetenzbereiche zeigen, dass auch in den Deutsch-Einzeltests die Mehrsprachigen und die nicht deutsch Erstsprachigen fast immer schlechter abschneiden als die deutsch Erstsprachigen. Für die differenziellen Leistungen der beiden Gruppen gilt: In allen Kompetenzbereichen sind die Mehrsprachigen den nicht deutsch Erstsprachigen überlegen. Interessanterweise schneiden die Mehrsprachigen in der Rechtschreibung sogar besser ab als die deutsch Erstsprachigen. Die Wortschatzleistungen der nicht deutsch Erstsprachigen fallen besonders niedrig aus und tragen in besonderem Maße zu den schwächeren Gesamtttestleistungen dieser Gruppe in den Deutschtests bei, dies gilt jedoch auch für die Gruppe der Mehrsprachigen. Die schwache Lesekompetenz der nicht deutsch Erstsprachigen spiegelt die Ergebnisse der PISA 2003 Studie wider, in der allerdings der Migrationsstatus als Kriterium zur Gruppenbildung zu Grunde gelegt wurde und ein völlig anderer Test zur Messung der Kompetenz im Deutschen zur Anwendung kam. Die Punktedifferenzen sind also demnach nicht vergleichbar (OECD 2006). Das Lesen, die Argumentation und die Pragmatikkomponente der Textproduktion

sind Kompetenzbereiche, in denen sich die Mehrsprachigen nicht signifikant von denen der deutsch Erstsprachigen unterscheiden.

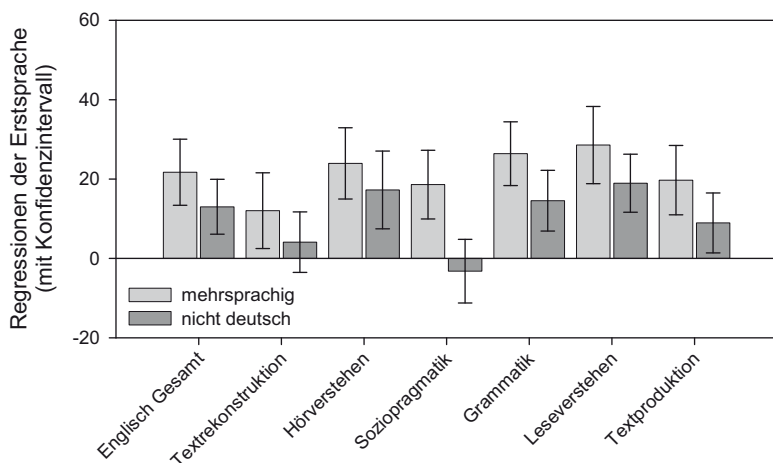


Abbildung 20.5: Effekte des sprachlichen Hintergrundes auf Kompetenzbereiche im Englischen zum Ende des neunten Jahrgangs unter Kontrolle von Bildungsgang, sozioökonomischem Hintergrund, kognitiver Grundfähigkeit und Geschlecht: Kontraste der Leistungen von Mehrsprachigen und nicht deutsch Erstsprachigen gegenüber deutsch Erstsprachigen. Anm.: Bewusstheit Soziopragmatik wurde aus technischen Gründen zu Beginn des Schuljahres erfasst.

Betrachtet man nun die Teilleistungsbereiche im Englischen, so stellt man signifikante Unterschiede zum Ende der neunten Jahrgangsstufe fest, die zum großen Teil denen der Englisch Gesamtleistung vergleichbar sind. Prägnant treten die hohen Testleistungen der Mehrsprachigen beim Leseverstehen, bei der Komponente Grammatik, der Sprachbewusstheit im Englischen und beim Hörverstehen hervor. In diesen Kompetenzbereichen zeigen auch die nicht deutsch Erstsprachigen deutlich bessere Testleistungen als die deutsch Erstsprachigen, sobald die konfundierenden Variablen kontrolliert werden. Die Leistungen der nicht deutsch Erstsprachigen liegen nur im C-Test (nicht signifikant besser) und in Soziopragmatik (nicht signifikant schlechter) in etwa auf dem Niveau der deutsch Erstsprachigen, in den anderen Kompetenzbereichen liegen sie deutlich darüber.

Die Testleistungen in der Komponente Soziopragmatik der Sprachbewusstheit im Englischen unterscheiden sich auffällig zwischen Mehrsprachigen und nicht deutsch Erstsprachigen. Der DESI-Test zur Soziopragmatik im Englischen bildet den angemessenen Gebrauch der Fremdsprache in Bezug auf den Kontakt, die Situation und den Adressaten ab. Die Voraussetzung für einen angemessenen Sprachgebrauch ist die Fähigkeit, Sprechintentionen erkennen und gedanklich einordnen zu können. Mehrsprachigen scheint dieser Teil des sozialen und interkulturellen Rollenverhaltens besser zu gelingen als monolingual Erstsprachigen. Interessanterweise lässt sich ein Zusammenhang zwischen der kulturellen Angemessenheit der Sprache, so wie sie im Englischtest zur Soziopragmatik erfasst wird, und interkultureller Kompetenz fest-

stellen (vgl. Kapitel 17). Der ethnorelative Typus der interkulturellen Kompetenz zeichnet sich dadurch aus, dass Schülerinnen und Schüler Bewusstheit und Akzeptanz gegenüber kultureller Verschiedenheiten entwickelt und kulturelle Fähigkeiten zur Anpassung an anderskulturelle Situationen ausgebildet haben. Der im Kapitel 17 berichtete Befund, dass Mehrsprachige vermehrt dem ethnorelativen Typus der interkulturellen Kompetenz zuzuordnen sind, kann damit als Unterstützung der Annahme gewertet werden, dass sich frühe Mehrsprachigkeit sowohl in Bezug auf den sozio-pragmatischen Aspekt des Fremdsprachengebrauchs als auch auf den Erwerb einer ethnorelativen interkulturellen Orientierung positiv auswirkt.

20.2 Sprachvitalität

Der Erwerb der Erstsprache und der einer Zweitsprache werden als miteinander verbunden angesehen (Gogolin/Neumann/Roth 2003). Ein spezifischer Zusammenhang wird auf Grund theoretischer Überlegungen, wie beispielsweise der Interdependenzhypothese angenommen (Cummins 1979). Inwieweit die Herkunftssprachen aber auch empirisch nachweisbar als Ressource für den Erwerb des Deutschen und der ersten schulischen Fremdsprache (zumeist des Englischen) dienen, ist jedoch bislang eine weitgehend offene Forschungsfrage (Maas/Mehlem/Schroeder 2004). Trotz der im Verlauf der Diskussionen eingenommenen positiven Sicht der Mehrsprachigkeit blieb die Skepsis bestehen, Kinder könnten durch die Anforderungen einer mehrsprachigen Umwelt überfordert werden. Besonders bekannt wurde das Modell von Cummins (1993), der von einer „Schwelle“ ausgeht: Erst wenn die Erstsprache bis zu einem gewissen Kompetenzniveau erworben worden sei, könne eine weitere Sprache erfolgreich hinzugelehrt werden. Gesähe dies nicht, so würde sich die problematische Situation der „subtraktiven Zweisprachigkeit“ einstellen, bei der in keiner der erworbenen Sprachen ein Niveau erreicht werden könne, das dem von Muttersprachigen vergleichbar wäre (Cummins 1979). Dies entspricht der These von der „Halbsprachigkeit“ (Skutnabb-Kangas 1984). Solche Thesen, die auch im schulischen Alltag in der einen oder anderen Form vertreten werden, lassen sich zum Teil nur äußerst schwer in empirisch überprüfbare Hypothesen übertragen.

In Large-Scale-Assessments sind Tests der Erstsprachenkompetenz aus pragmatischen Gründen nicht durchführbar – für DESI hätten in mehr als 26 Sprachen Tests konstruiert werden müssen. Andere Studien, die sich mit dem Verlauf des Zweitspracherwerbs im Migrationskontext beschäftigen, nutzen Selbsteinschätzungen der Sprachkompetenz, um die Fertigkeiten in der Erstsprache abzubilden (Broeder/Extra 1999; Haug 2005). Diesen Beispielen folgend wurden in DESI von den „nicht deutsch Erstsprachigen“ und den „Mehrsprachigen“ Selbsteinschätzungen der Kompetenz in der nicht-deutschen Erstsprache erbeten. Auch wenn die Schwellentheorie damit nicht adäquat abgebildet werden kann, so müsste, der Argumentation der Interdependenzhypothese folgend, eine positive Wirkung der Erstsprachenkompetenz auf die sprachlichen Leistungen im Deutschen und im Englischen erkennbar sein.

Die Selbsteinschätzung erfolgte in zweierlei Hinsicht, zum einen sollten die Schülerinnen und Schüler angeben, wie gut sie Äußerungen in der Erstsprache verstehen können, zum anderen, wie gut sie sich darin ausdrücken können. Die Selbsteinschätzungen in den beiden Bereichen „Verstehen“ und „Sich-Ausdrücken“ beziehen sich sowohl auf die mündliche als auch auf die schriftliche Kommunikation. Eine weitere Differenzierung zwischen der Selbsteinschätzung des erstsprachlichen Hörverstehens beziehungsweise des Leseverstehens und der mündlichen Ausdrucksfähigkeit gegenüber der Schreibkompetenz ist anhand dieser Skalen nicht möglich. Die Selbsteinschätzungen wurden in Anlehnung an Kriterien des Gemeinsamen Europäischen Referenzrahmens für Sprachen (Trim/North 2001) erfasst. Mit Hilfe von kurzen Charakterisierungen auf der Grundlage der sechs Stufen des Referenzrahmens in Form von „Can-do-Statements“ und einer zusätzlichen Nullstufe wurden jeweils siebenstufige Skalen konstruiert.⁶ Die Schülerinnen und Schüler sollten jeweils das höchste Niveau ankreuzen, das sie ihrer Einschätzung nach erreichen.

Zweifelloos handelt es sich dabei um eine subjektive und gewissermaßen grobe Einschätzungen eigener Kompetenzen. In den Abbildungen 20.6 und 20.7 sind die Regressionskoeffizienten für die jeweiligen Selbsteinschätzungsskalen bei der Vorhersage der Testleistungen mit Konfidenzintervallen dargestellt. Die Regressionsanalysen basieren auf den Guppen der mehrsprachigen und nicht-deutschen Schülerinnen und Schüler. Als Kontrollvariablen wurden der Bildungsgang, der sozioökonomische Status, die kognitive Grundfähigkeit, das Geschlecht und der sprachliche Hintergrund (mehrsprachig vs. nicht-deutsch) einbezogen. Schülerinnen und Schüler, die sich in ihrer Erstsprache um ein Niveau kompetenter eingestuft haben als andere zeigen auf der vertikalen Achse eine entsprechend höhere oder niedrigere Leistungsausprägung. Es zeigt sich ein fast durchweg positiver Zusammenhang zwischen der Sprachkompetenzeinschätzung in der Herkunftssprache und den Deutschtestleistungen. Dieser wird zwischen dem sich in der Herkunftssprache Ausdrücken-Können und der Deutsch-Gesamtleistung sowie dem Teilleistungsbereich Textproduktion (Semantik und Pragmatik) und dem Argumentieren signifikant. Für den Kompetenzbereich Argumentation ist auch bereits die Einschätzung des Verstehen-Könnens in der Erstsprache ein signifikant positiver Prädiktor.

6 In Bezug auf das Verstehen waren die Stufen gekennzeichnet durch „verstehe kaum etwas / nur ganz einfache Sätze / einfache Sätze und häufig gebrauchte Ausdrücke / das Wichtigste, wenn es um vertraute Dinge geht / Hauptinhalte auch bei schwierigen Texten / die meisten schwierigen Texte, auch längere / alle schwierigen Texte mühelos“; in Bezug auf das Sich-Ausdrücken können „kann mich kaum verständigen / auf eine einfache Art verständigen, wenn Gesprächspartner langsam und deutlich sprechen und helfen / in Situationen verständigen, in denen es um bekannte Dinge geht / zusammenhängend über vertraute Dinge unterhalten / zu den meisten Themen klar und genau ausdrücken / bei schwierigen Themen klar, strukturiert, ausführlich äußern ohne öfter nach Wörtern suchen zu müssen / spontan, sehr flüssig und genau ausdrücken, bei schwierigen Themen feinere Bedeutungsunterschiede deutlich machen“.

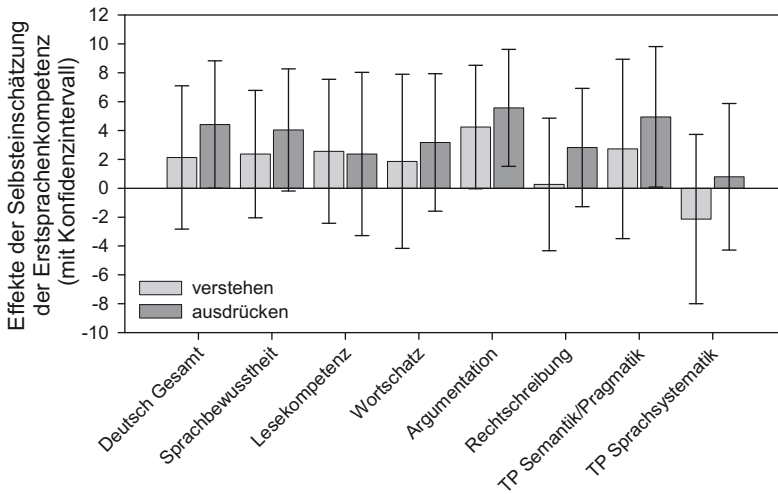


Abbildung 20.6: Effekte der Selbsteinschätzung der Erstsprachenkompetenz im Verstehen- und Sich-Ausdrücken-Können zum Ende des neunten Jahrgangs auf die Kompetenzbereiche im Deutschen unter Kontrolle von Bildungsgang, sozioökonomischem Hintergrund, kognitiver Grundfähigkeit, Geschlecht und Mehr- bzw. Einsprachigkeit für die Gruppe der Schüler, die (auch) eine andere Sprache als Deutsch als Erstsprache erworben haben.

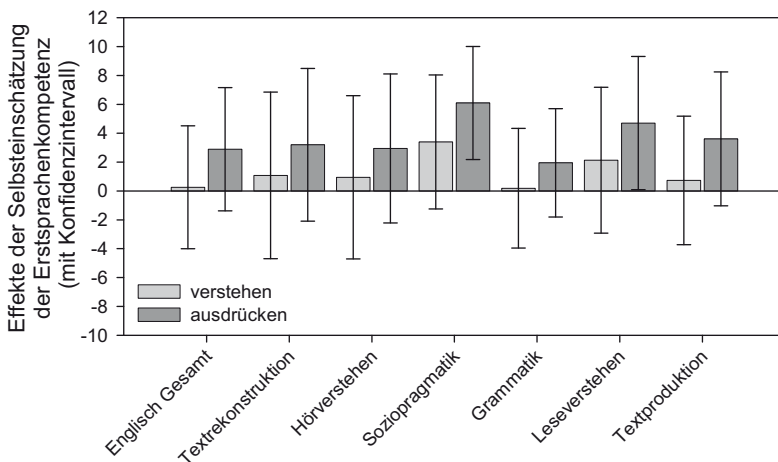


Abbildung 20.7: Effekte der Selbsteinschätzung der Erstsprachenkompetenz im Verstehen- und Sich-Ausdrücken-Können zum Ende des neunten Jahrgangs auf die Kompetenzbereiche im Englischen unter Kontrolle von Bildungsgang, sozioökonomischem Hintergrund, kognitiver Grundfähigkeit, Geschlecht und Mehr- bzw. Einsprachigkeit für die Gruppe der Schüler, die (auch) eine andere Sprache als Deutsch als Erstsprache erworben haben. Anm.: Bewusstheit Soziopragmatik wurde aus technischen Gründen zu Beginn des Schuljahres erfasst.

Für die Kompetenzen im Englischen zeigen sich ebenfalls durchgehend positive Effekte der Herkunftssprachenkompetenz. Die einzige statistisch bedeutsame Regression konnte jedoch nur für die Teilkomponente „Soziopragmatik“ im Englischen gefunden werden. Wenn sich die Schülerinnen und Schüler in ihrer Erstsprache gut ausdrücken können, dann geht dies offenbar mit einer höheren soziopragmatischen Leistung im Englischen einher.

Die selbsteingeschätzte kommunikative erstsprachliche Kompetenz („Sich-Ausdrücken-Können“) indiziert offenbar eine für den Erwerb der deutschen Sprache förderliche Bedingung für Schülerinnen und Schüler nicht deutscher Erstsprache und für Mehrsprachige. Das Argumentieren im Deutschen – ein kommunikativer Kompetenzbereich – scheint denjenigen leichter zu fallen, die ihre Erstsprache gut verstehen und sich gut in ihr ausdrücken können. Die Angemessenheit des sprachlichen Ausdrucks, so wie sie in der Textproduktion in Bezug auf Semantik und Pragmatik erfasst wurde, wird augenscheinlich ebenfalls durch die kommunikative erstsprachliche Kompetenz gefördert. Dies gilt ebenso für den soziopragmatischen Kompetenzbereich im Englischen. Aus diesen Ergebnissen lässt sich die Vermutung ableiten, dass sich gute erstsprachliche Kompetenzen beim Erwerb weiterer Sprachen auf kommunikative und pragmatische Sprachkompetenzen niederschlagen.

Soweit kommunikative Aspekte betrachtet werden, widersprechen die Ergebnisse der Interdependenzhypothese, also der Annahme einer Abhängigkeit von erstsprachlicher Kompetenz und Zweit- oder Fremdsprachenkompetenzen, nicht. Dass es keine signifikanten Effekte der Selbsteinschätzung der Erstsprachkompetenz auf die anderen in DESI erhobenen Deutsch- und Englischkompetenzen gibt, kann mehrere Gründe haben. Zum einen können die nicht weiter differenzierten Selbsteinschätzungsskalen zu diesem Ergebnis geführt haben. Da es nicht möglich war, mündliche und schriftliche Kompetenzen differenziert abzubilden, haben die Schülerinnen und Schüler unter Umständen vor allem über ihre mündlichen Fähigkeiten in ihrer Erstsprache Auskunft gegeben. Hierfür spricht auch die Formulierung der „Can-do-Statements“ der Skalen. Geht man davon aus, dass die Erstsprache vor allem die Sprache der privaten und familiären Kommunikation ist (Esser 2001), wird Mündlichkeit hier eine besonders große Rolle spielen. Es erscheint daher plausibel, dass sich diese berichtete mündliche Beherrschung der Erstsprache vor allem auf kommunikative Aspekte der Zweit- und Drittsprache auswirkt.

Dennoch kann die Interdependenzhypothese in ihrer Allgemeinheit in Zweifel gezogen werden. Diese Zweifel sind durch die gegenwärtig berichtete unsichere Befundlage zur Auswirkung der Erstsprachenkompetenz auf den Zweit- und Drittspracherwerb begründbar (zusammenfassend: Cenoz 2003; Hopf 2005). Neben positiv gerichteten Korrelationen (Sanz 2000) werden auch neutrale oder negative Zusammenhänge (van Gelderen/Schoonen/de Glopper 2003) berichtet. Diese Studien fanden nicht nur in unterschiedlichen Ländern (Spanien versus Niederlande) statt und spiegeln dadurch eine unterschiedliche Situation der Mehrsprachigen wider (mehrsprachige Nation versus Migrationskontext), sondern sie beziehen sich auch auf unterschiedliche Teilaspekte sprachlicher Kompetenz. Ein allumfassendes

Urteil über die Interdependenzhypothese können diese Studien ebenso wenig fällen, wie dies die Ergebnisse DESI-Studie können. Festzuhalten bleibt jedoch, dass sich im Rahmen der DESI-Untersuchung keine negativen Effekte der selbstberichteten Erstsprachenkompetenz auf die Kompetenz in der Zweit- und Drittsprache gezeigt haben.

Aus den vorliegenden korrelativen Daten der DESI-Studie lassen sich keine Aussagen darüber ableiten, inwieweit die Förderung der Erstsprache für den Erwerb des Deutschen und des Englischen vorteilhaft wäre. Hierzu wären Längsschnittstudien und experimentelle Vorgehensweisen notwendig. Insbesondere aber bedarf es kompetenzorientierter psychometrisch fundierter Testverfahren, welche die Erstsprachenkompetenz möglichst objektiv zu erfassen in der Lage sind. Erst dann lassen sich die skizzierten theoretischen Erwartungen empirisch überprüfen und Praxisempfehlungen äußern.

20.3 Effekte der Zusammensetzung der Schülerschaft auf die sprachlichen Leistungen

Die Lernleistungen in einer Schulklasse hängen nicht nur von den individuellen Voraussetzungen der Schülerinnen und Schüler ab, sondern werden auch von der Zusammensetzung der Schülerschaft beeinflusst. Eine ganze Reihe von Schulleistungsstudien in verschiedenen Ländern ergab, dass hohe Anteile von Schülerinnen und Schülern aus Familien mit niedrigem sozioökonomischem Status und von Migranten in der Schulklasse mit niedrigeren Testleistungen einhergehen (z.B. OECD 2004). Darüber hinaus stellte sich für Deutschland der Einfluss des Bildungsgangs und der mittleren kognitiven Fähigkeiten der Schülerschaft als besonders einflussreiche Kompositionseffekte heraus (Baumert/Stanat/Waterman 2006).

Wenn Mehrsprachige bessere Leistungen im Fremdsprachenunterricht zeigen, dann wäre zu erwarten, dass sich die Richtung der bislang ermittelten Kompositionseffekte umkehren sollte: Ein relativ hoher Anteil von mehrsprachigen Schülerinnen und Schülern in einer Klasse sollte dann die Gesamtestleistungen im Englischen erhöhen.

Für die Prüfung von Kompositionseffekten bieten sich Zweiebenenmodelle an. Die generelle Fragestellung ist, ob die Zusammensetzung der Schülerschaft über die individuellen Effekte der Schülermerkmale hinaus einen Erklärungsgewinn für das Zustandekommen der Testleistungen erbringt. Da auch auf der Ebene der Schulklasse Konfundierungen vorliegen, ist es zweckmäßig, mit Hilfe gestaffelter Analysemodelle zu versuchen, die Effekte des Ausgangsmodells durch Effekte von hinzugefügten Variablen in Folgemodellen zu „erklären“. Dies ist dann der Fall, wenn sich die Regressionskoeffizienten verringern.

Analysen dieser Art wurden von beispielsweise von Rüesch (1998) und Stanat (2006) durchgeführt. Rüesch berichtet über das Verschwinden des Einflusses des Migrantenanteils in der Klasse, wenn Merkmale der sozialen Zusammensetzung mehrbenenanalytisch kontrolliert werden. In Bezug auf die Analyse der PISA-Daten

stellte Stanat hingegen nur eine Reduktion, aber kein Verschwinden der Einflussgröße des Migrantenanteils fest. Erst bei zusätzlicher Kontrolle der leistungsbezogenen Zusammensetzung der Schülerschaft wurde der Effekt des Migrantenanteils nicht mehr signifikant. Die Wirkungen eines hohen Migrantenanteils in der Klasse lassen sich demnach durch den sozialen Hintergrund und die kognitiven Grundfähigkeiten der Schülerinnen und Schüler „erklären“.

Ausgehend von den bisher berichteten Ergebnissen wird im Folgenden geprüft, inwieweit sich die Zusammensetzung der Schülerschaft hinsichtlich Erstsprache, sozioökonomischem Hintergrund, kognitiven Grundfähigkeiten und Geschlecht in den einzelnen Bildungsgängen über die Effekte der individuellen Merkmale hinaus auf die Gesamtleistungen in Deutsch und Englisch auswirkt. Entsprechend wird eine Reihe von Modellen konstruiert, welche die genannten Effekte schrittweise abzuschätzen erlaubt.

In der Tabelle 20.6 sind die Ergebnisse aus Zweiebenenanalysen für die Gesamtleistungen im Deutschen und in Tabelle 20.7 für die Gesamtleistungen im Englischen auf der Grundlage von fünf hierarchisch ineinander verschachtelten Modellen zusammengefasst. Das Modell 1 bezieht lediglich individuelle Schülermerkmale mit ein und beschränkt sich somit ausschließlich auf die Schülerebene. Vergleichbar mit den Ergebnissen der Analysen von Baumert, Waterman und Stanat (2006) stellen die kognitiven Grundfähigkeiten auf dieser Ebene die stärkste Bedingung für die Leistungen in beiden Sprachen dar. Der Anstieg von einer Standardabweichung bei den Ergebnissen des KFT erhöht die sprachlichen Leistungen um 51 Punkte im Deutschen bzw. 49 Punkte im Englischen. Was das Fach Deutsch betrifft, so ist der Effekt der nicht deutschen Erstsprachigkeit ebenso deutlich: Jugendliche, die ausschließlich eine andere Sprache als Deutsch als Erstsprache erworben haben erzielen 52 Punkte weniger als ihre Mitschüler mit Deutsch als Erstsprache. Bei den Mehrsprachigen ist dieser negative Effekt mit minus 12 Punkten weit geringer. Deutlich verschieden sind hierzu die Ergebnisse im Englischen. Analog zu den Ausführungen im Abschnitt 20.1 findet man einen positiven Effekt der Mehrsprachigkeit (24 Punkte) und einen, wenn auch nicht signifikanten, positiven Effekt auf Grund einer nicht deutschen Erstsprachenbiographie. Der sozioökonomische Status prädiziert, wie in den Analysen der anderen Kapitel bereits gezeigt, die sprachlichen Leistungen deutlich; weiterhin haben Mädchen bei den Sprachkompetenztestungen besser abgeschnitten als männliche Jugendliche, besonders in den Deutschleistungen. Soweit die Betrachtung auf der Individualebene.

Das Modell 2 bezieht zusätzlich auf der Schulklassenebene die Anteile der nicht deutsch Erstsprachigen und der Mehrsprachigen ein. Steigt der prozentuale Anteil der nicht deutsch Erstsprachigen in der Klasse, dann sinken die Leistungen sowohl im Deutschen als auch im Englischen. Andererseits erhöhen sich die Leistungen, wenn der Anteil der Mehrsprachigen in der Klasse zunimmt. Die Beträge sind zwar

klein, aber mit Ausnahme der Deutschleistungen der Mehrsprachigen signifikant⁷. Im Modell 2 werden Unterschiede zwischen den Bildungsgängen nicht modelliert. Wie Tabelle 20.3 zeigt, sind die nicht deutsch Erstsprachigen jedoch seltener im Gymnasium und in der Realschule anzutreffen. Um diese Konfundierung zu berücksichtigen wird das Modell 3 gegenübergestellt.

Modell 3 berücksichtigt auch den Bildungsgang. Die Effektgrößen des Modells 2 auf Grund der Anteile nicht deutsch Erstsprachiger und der Mehrsprachigen reduzieren sich deutlich zugunsten der Bildungsgangeffekte. Die Tendenz jedoch bleibt erhalten: Der Anteil der nicht deutsch Erstsprachigen verringert die Leistungen im Deutschen, während er sich auf die Leistungen im Englischen nicht signifikant auswirkt. Der Anteil der Mehrsprachigen wirkt sich positiv auf die Leistungen aus, im Englischen mehr als im Deutschen. Die Koeffizienten auf der Individualebene sind mit denen des Modells 2 weitgehend vergleichbar.

Das Modell 4 bezieht die sozioökonomische Zusammensetzung der Schülerschaft mit ein. Der Effekt des mittleren sozioökonomischen Status in der Klasse ist deutlich, im Englischen mehr als im Deutschen.

Nimmt man die mittlere kognitive Grundfähigkeit als einen weiteren Faktor der Zusammensetzung der Schülerschaft in das Modell auf (Modell 5), dann verändern sich zwar die Koeffizienten für die Bildungsgänge und den sozioökonomischen Status, nicht aber die Kompositionseffekte der Erstsprache. Beide Anteile spielen in Bezug auf Deutsch keine Rolle. Die Englischleistungen in der Klasse sind signifikant besser, wenn mehr Schüler nicht deutscher Erstsprache und Mehrsprachige anwesend sind, sofern die sozioökonomischen Voraussetzungen und der Bildungsgang kontrolliert werden.

In Entsprechung zu den Analysen von Rüesch und Stanat verschwindet bei den Deutschleistungen nach Berücksichtigung des Bildungsgangs, der sozialen Herkunft und der kognitiven Grundfähigkeiten auf Klassenebene der Effekt der Zusammensetzung der Schülerschaft im Hinblick auf den Anteil der ausschließlich nicht deutsch Erstsprachigen. Interessanterweise erfolgt im Englischen eine Vorzeichenumkehr: Bei entsprechenden statistischen Kontrollen ist ein positiver Effekt auf die Englischgesamtleistung von einem hohen Anteil nicht deutsch Erstsprachiger festzustellen.

Durch den zuvor geschilderten Einfluss der Klassenzusammensetzung erfährt der Einfluss der kognitiven Grundfähigkeiten- und sozioökonomischen Status-Koeffizienten auf der individuellen Ebene eine deutliche Reduktion. Ebenso sinkt die Vorhersagekraft der Geschlechtszugehörigkeit. Die individuellen Effekte der Erstsprache verändern sich ebenfalls, jedoch die Richtung der Einflüsse bleibt erhalten.

7 Die geringe Größe der Koeffizienten muss im Zusammenhang mit den prozentualen Anteilen gesehen werden. Bei einer durchschnittlichen Klassengröße von 20 Schülern macht jeder Schüler 5% aus. Erhöht sich die Anzahl der Mehrsprachigen um einen Schüler in der Klasse, dann steigt die mittlere Englischleistung in der Klasse um $5 \text{ mal } 2,1 = 10,5$ Punkte.

Die in den Tabellen 20.6 und 20.7 dargestellten Analysen bestätigen die starke Abhängigkeit der Schülerleistungen vom Bildungsgang und dem sozialen Hintergrund. Beide Einflussgrößen sind jedoch stark mit dem Indikator Erstsprache verknüpft. Kontrolliert man sie statistisch, so wird die sprachliche Zusammensetzung der Schülerschaft in Bezug auf Deutsch unbedeutend. Für die Englischleistungen hingegen ist die Anwesenheit von nicht deutsch Erstsprachigen und Mehrsprachigen förderlich.

Bei Betrachtung des Einflusses der Zusammensetzung der Schülerschaft wird die Folgerung unterstützt: Mehrsprachigkeit ist nicht nur individuell eine günstige Voraussetzung für das Lernen einer Fremdsprache in der Schule, sie trägt auch im Klassenverband zur Leistungssteigerung aller bei.

Tabelle 20.6 Effekte der Zusammensetzung der Schülerschaft auf die Deutschleistungen. Standardisierte Koeffizienten sind kursiv gedruckt.

Prädiktor	Modell 1 ³	Modell 2 ⁴	Modell 3 ⁴	Modell 4 ⁴	Modell 5 ⁴
Schülerebene					
Kognitive Grundfähigkeiten ¹	50.5* 0.51*	23.4* 0.38*	22.9* 0.37*	22.8* 0.37*	21.7* 0.36*
Sozioökonomischer Status ¹	22.5* 0.22*	5.0* 0.08*	4.6* 0.07*	4.0* 0.06*	4.0* 0.06*
Schülerin (=1, männl.=0)	37.7* 0.19*	27.7* 0.22*	27.7* 0.22*	27.7* 0.22*	27.7* 0.23*
Erstsprache nicht Deutsch (=1, Rest=0)	-51.7* -0.17*	-43.7* -0.23*	-44.0* -0.23*	-44.2* -0.23*	-44.4* -0.24*
Mehrsprachig (=1, Rest=0)	-11.8* -0.03*	-17.0* -0.06*	-17.1* -0.06*	-17.0* -0.06*	-17.2* -0.06*
Klassenebene					
% Anteil Schüler, deren Erstsprache nicht Deutsch ist		-1.8* -0.42*	-0.5* -0.11*	-0.1 -0.03	0.1 0.01
% Anteil Schüler, die mehr- sprachig aufgewachsen sind		1.1 0.11	0.6 0.06	0.6 0.06	0.5 0.05
Realschule (=1, Rest=0)			59.9* 0.43*	46.0* 0.33*	25.9* 0.18*
Gymnasium (=1, Rest=0)			131.8* 0.88*	81.0* 0.53*	45.5* 0.30*
Mittlerer sozioökonomischer Status ²				58.7* 0.42*	43.5* 0.31*
Mittlere kognitive Grundfähigkeiten ²					41.0* 0.41*

Anmerkungen: * $p < .05$; ¹z-transformiert; ²auf der Basis von z-transformierten Werten schulklassenweise aggregiert; ³berechnet mit WESVAR; ⁴berechnet mit Mplus

Tabelle 20.7 Effekte der Zusammensetzung der Schülerschaft auf die Englischleistungen.

Prädiktor					
	Modell 1 ³	Modell 2 ⁴	Modell 3 ⁴	Modell 4 ⁴	Modell 5 ⁴
Schülerebene					
Kognitive Grundfähigkeiten ¹	49.3* 0.49*	17.3* 0.34*	16.8* 0.33*	16.7* 0.33*	16.1* 0.32*
Sozioökonomischer Status ¹	26.9* 0.26*	3.7* 0.07*	3.5* 0.07*	3.0* 0.06*	3.0* 0.06*
Schülerin (=1, männl.=0)	27.4* 0.14*	15.2* 0.15*	15.2* 0.15*	15.2* 0.15*	15.2* 0.15*
Erstsprache nicht Deutsch (=1, Rest=0)	6.3 0.02	4.5 0.03	4.3 0.03	4.2 0.03	4.1 0.03
Mehrsprachig (=1, Rest=0)	23.8* 0.05*	8.2* 0.04*	8.1* 0.04*	8.2* 0.04*	8.1* 0.04*
Klassenebene					
% Anteil Schüler, deren Erstsprache nicht Deutsch ist		-1.7* -0.33*	-0.1 -0.01	0.4* 0.07*	0.5* 0.09*
% Anteil Schüler, die mehr- sprachig aufgewachsen sind		2.1* 0.17*	1.5* 0.12*	1.5* 0.12*	1.5* 0.12*
Realschule (=1, Rest=0)			66.0* 0.40*	48.8* 0.29*	33.0* 0.20*
Gymnasium (=1, Rest=0)			168.3* 0.93*	105.6* 0.58*	77.6* 0.43*
Mittlerer sozioökonomischer Status ²				72.1* 0.44*	60.1* 0.36*
Mittlere kognitive Grundfähigkeiten ²					32.0* 0.27*

Anmerkungen: * $p < .05$; ¹z-transformiert; ²auf der Basis von z-transformierten Werten schulklassenweise aggregiert; ³berechnet mit WESVAR; ⁴berechnet mit Mplus

20.4 Schlussfolgerungen

Die eingangs gestellte Frage, inwiefern sich der in bisherigen Large-Scale-Assessments für Deutschland beobachtete starke Zusammenhang von Migrationsstatus, sozialer Herkunft und Schülerleistungen auch bei sprachlichen Leistungen widerspiegelt, ist durch die Aufteilung der Schülerschaft in deutsch oder nicht deutsch Erstsprachige und Mehrsprachige differenziert zu beantworten. Bei DESI stehen dabei nicht die vermuteten Wirkungen des Migrationsstatus, sondern der ersten erworbenen Sprache oder Sprachen im Vordergrund. Mehrsprachige weisen weit bessere Deutschleistungen auf als nicht deutsch Erstsprachige. Sie erzielen im Englischen im Verhältnis zu den deutsch Erstsprachigen vergleichbare, in der Realschule sogar bessere Resultate. Unter statistischer Kontrolle von Bildungsgang, sozialer Herkunft, kognitiven Grundfähigkeiten und Geschlecht ergibt sich für die Deutsch- und Englischgesamtleistungen ein gegenläufiges Bild. Im Vergleich zu den deutsch Erstsprachigen sind die Deutschleistungen der nicht deutsch Erstsprachigen erheblich

niedriger, die der Mehrsprachigen immer noch deutlich niedriger, aber nicht ganz so weit abgeschlagen. Dies verhält sich im Englischen umgekehrt: Die Mehrsprachigen erzielen deutlich bessere, die nicht deutsch Erstsprachler etwas bessere Leistungen als die deutsch Erstsprachigen.

Die Ergebnisse legen den Schluss nahe, dass mehrsprachige Lernumwelten offensichtlich günstige Bedingungen für die Leistungen im Fach Englisch darstellen. Beim Erwerb von Fremdsprachen scheint die Schülerschaft, die (auch) eine andere Erstsprache als Deutsch erworben hat, über ein Potenzial zu verfügen, das es ihr ermöglicht, überlegene Leistungen in den Leistungstest im Englischen zu erzielen. Falls es sich dabei um ein generelles Sprachlernpotenzial handeln sollte, dann kann dieses Potenzial im Deutschunterricht offenbar nicht in vergleichbarer Weise gut genutzt werden. Die Beherrschung der Unterrichtssprache Deutsch wird jedoch als Voraussetzung für den schulischen und beruflichen Erfolg (Baumert/Schümer 2001) und die „Integration“ mit der deutschen Gesellschaft (oder gar „Assimilation“: Esser 2001) angesehen. Vermutlich wirkt sich die Art und Weise, wie im Deutschunterricht – anders als im Englischunterricht – Sprache vermittelt wird, auf die jeweiligen Leistungen dieser Schülerschaft aus: im Englischunterricht steht die Sprache selbst viel mehr im Vordergrund. Wenn dem so ist, dann wäre ein auf die Sprache konzentrierter Deutsch- und Sachfachunterricht insbesondere für die Schülerinnen und Schüler nicht deutscher Erstsprache und für die Mehrsprachigen zur besseren Entwicklung ihrer Deutschkompetenzen hilfreich.

Der viel diskutierte Zusammenhang zwischen der Sprachvitalität in der Erstsprache und dem Erwerb weiterer Sprachen kann durch die DESI-Studie auf der Grundlage von Selbsteinschätzungen erstsprachlicher Kompetenzen empirisch beleuchtet werden. Negative Einflüsse der Sprachvitalität in der Herkunftssprache auf die Teilkompetenzen in der Zweit- und Drittsprache können nicht nachgewiesen werden. In Bezug auf kommunikative und pragmatische Aspekte zeigen sich positive Zusammenhänge mit der Herkunftssprachenkompetenz. Auch die Deutsch Gesamtestleistung ist erhöht, wenn die Schülerinnen und Schüler der Überzeugung sind, sie könnten sich in ihrer Erstsprache gut ausdrücken. Die Frage aber, ob die gezielte Förderung in der Herkunftssprache die sprachlichen Leistungen im Deutschen und Englischen allgemein verbessern würde, lässt sich auf Grund der hier zu Grunde gelegten Daten nicht beantworten.

Mithilfe von Zweiebenenmodellen, in denen der Bildungsgang, die soziale Herkunft, die kognitiven Grundfähigkeiten, das Geschlecht und die Erstsprachen statistisch kontrolliert werden, lässt sich zeigen, dass die Zusammensetzung der Schülerschaft in Bezug auf die Erstsprachen für die Leistungen im Deutschen unbedeutend ist. Im Englischen dagegen erweist sich die Anwesenheit sowohl von nicht deutsch Erstsprachigen als auch von Mehrsprachigen als leistungssteigernd für die ganze Schulklasse. Das bedeutet, dass die Schülerinnen und Schüler, die über mehrsprachige Kompetenzen verfügen, ein Potenzial in Bezug auf das Fremdsprachenlernen für den Klassenverband darstellen, das von den Lehrpersonen ausgeschöpft werden sollte.

Ingesamt gesehen scheinen insbesondere Mehrsprachige, aber auch nicht deutsch Erstsprachige von ihrer mehrsprachigen Lernumwelt beim Erwerb einer schulischen Fremdsprache zu profitieren. Darüber hinaus sind diese Schülerinnen und Schüler in der Lage, zu einer Leistungssteigerung ihrer Schulklasse beitragen zu können. In welcher Weise diese positive Beeinflussung innerhalb des Unterrichts stattfindet, kann vor dem Hintergrund der vorliegenden Analyse nicht beantwortet werden, dies ist jedoch eine Frage, der im Rahmen von Analysen der DESI-Unterrichtsvideographien nachgegangen wird.

Literatur

- Baumert, J./Schümer, G. (2001): Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich, S. 323-407.
- Baumert, J./Stanat, P./Watermann, R. (2006): Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bialystok, E. (1991): Introduction. In: Bialystok, E. (Hrsg.): Language processing in bilingual children. Cambridge: Cambridge University Press, S. 1-9.
- Bialystok, E. (2005): Consequences of bilingualism for cognitive development. In: Kroll, J. F. / de Groot, A. M. B. (Hrsg.): Handbook of bilingualism. Psycholinguistic approaches: Oxford: Oxford Uni Press, S. 417-432.
- Broeder, P./Extra, G. (1999): Language, ethnicity & education. Clevedon: Multilingual matters.
- Cenoz, J. (2003): The additive effect of bilingualism on third language acquisition: A review. In: The International Journal of Bilingualism 7, H. 1, S. 43-69.
- Cenoz, J./Hufeisen, B./Jessner, U. (2001): Cross-linguistic influence in third language acquisition: Clevedon: Multilingual Matters.
- Cummins, J. (1979): Cognitive/academic language proficiency and linguistic interdependence, the Optimum Age question and some other matters. In: Working Papers on Bilingualism 19, S. 197-205.
- Cummins, J. (1993): Bilingualism and second language learning. In: Annual Review of Applied Linguistics 13, S. 51-70.
- Esser, H. (2001): **Integration und ethnische Schichtung**. Mannheimer Zentrum für Europäische Sozialforschung (MZES), Arbeitsbericht, Nr. 40. Mannheim.
- Gelderen, van A./ Schoonen, R./ Glopper, de K. (2003): Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension: A structural equation modelling approach. In: The International Journal of Bilingualism 7, H. 1, S. 7-25.
- Gogolin, I./Neumann, U./Roth, H.-J. (2003): Förderung von Kindern und Jugendlichen mit Migrationshintergrund (Gutachten). In: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (Hrsg.): Materialien zur Bildungsplanung und zur Forschungsförderung 107. Bonn, S. 1-19 und S. 36-60.
- Haug, S. (2005): Zum Verlauf des Zweitspracherwerbs im Migrationskontext. Eine Analyse der Ausländer, Aussiedler und Zuwanderer im Sozio-ökonomischen Panel. In: Zeitschrift für Erziehungswissenschaft 8, S. 263-284.
- Hopf, D. (2005): Zweisprachigkeit und Schulleistung bei Migrantenkindern. In: Zeitschrift für Pädagogik 51, H. 2, S. 236-251.
- Kielhöfer, B./Jonekeit, S. (1983): Zweisprachige Kindererziehung. Tübingen: Stauffenberg Verlag.

- Kommission der Europäischen Gemeinschaften (2005): Mitteilung der Kommission an das europäische Parlament und den Rat: Europäischer Indikator für Sprachenkompetenz. Brüssel: Kommission der Europäischen Gemeinschaften.
- Konsortium Bildungsberichterstattung (2006): Bildung in Deutschland: ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. Im Auftrag der Ständigen Konferenz der Kultusminister der Länder in der BRD und des Bundesministeriums für Bildung und Forschung. Bielefeld: Bertelsmann.
- Maas, U./Mehlem, U./Schroeder, C. (2004): Mehrsprachigkeit und Mehrschriftigkeit bei Einwanderern in Deutschland. In: Bade, K. / Bommes, M. / Münz, R. (Hrsg.): Migrationsreport 2004. Frankfurt am Main, S. 117-150.
- Mohanty, A. K./Perregaux, C. (1997): Language acquisition and bilingualism. In: Berry, J. W./ Dasen, P. R./Saraswathi, T. S. (Hrsg.): Handbook of cross-cultural psychology (2nd ed., Vol. 2: Basic Processes and Human Development) Boston, MA: Allyn & Bacon, S. 217-253.
- Neal, L. I./McCray, A. D./Webb-Johnson, G./Bridgest, S. T. (2003): The effects of African American movement styles on teachers' perceptions and reactions. In: The Journal of Special Education 37, S. 49-57.
- OECD (2003): Literacy skills for the world of tomorrow - Further results from PISA 2000. Paris: OECD.
- OECD (2006): Where immigrant students succeed. A comparative review of performance and engagement in PISA 2003. Paris: OECD Publishing.
- Peal, E./Lambert, W. (1962): The relation of bilingualism to intelligence. In: Psychological Monographs 76, S. 1-23.
- Rüesch, P. (1998): Spielt die Schule eine Rolle? Schulische Bedingungen ungleicher Bildungschancen von Immigrantenkindern – eine Mehrebenenanalyse. Frankfurt a.M.: Peter Lang.
- Sanz, C. (2000): Bilingual education enhances third language acquisition: Evidence from Catalonia. In: Applied Psycholinguistics 21, H. 1, S. 23-44.
- Segalowitz, N. S. (1981): Issues in the cross-cultural study of bilingual development. In: Triandis, H. C./Heron A. (Hrsg.): Handbook of cross-cultural psychology. Developmental Psychology 4. Boston, MA: Allyn & Bacon, S. 55-92.
- Skutnabb-Kangas, T. (1984): Bilingualism or not. Clevedon: **Multilingual Matters**.
- Stanat, P. (2006): Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft. In: Baumert, J./Stanat, P./Watermann, R. (Hrsg.): Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 51-80.
- Steinmüller, U. (1981): Begriffsbildung und Zweitspracherwerb. Ein Argument für den muttersprachlichen Unterricht. In: Essinger, H./ Hellmich, A. /Hoff, G. (Hrsg.): Ausländerkinder im Konflikt. Königstein/Ts.: Scriptor, S. 83-97.
- Trim, J. L. M./North, B. (2001): Gemeinsamer Europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. **Berlin: Langenscheidt**.
- United Nations (2007): Implementation of general assembly resolution 60/251 of 15 march 2006 entitled "human rights council". Report of the Special Rapporteur on the right to education, Vernor Muñoz. Addendum: Mission to Germany (13-21 February 2006). URL: <http://www.ohchr.org/english/bodies/hrcouncil/docs/4session/A.HRC.4.17.pdf> (Zugriff 04.11.2007)
- Weisgerber, L. (1966): Vorteile und Gefahren der Zweisprachigkeit. In: Wirkendes Wort 2, S. 73-89.

*Wolfgang Wagner / Andreas Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Eichler / Günther Thomé / Heiner Willenberg*

21 Selbstkonzept und Motivation im Fach Deutsch

21.1 Einführung

In der Motivationsforschung werden in der Tradition der sogenannten Erwartungs-mal-Wert-Theorien seit langem zwei sich ergänzende Konzepte unterschieden: Die Steuerung des Verhaltens durch den Anreizwert der Handlung selbst bzw. durch die damit verbundenen Folgen (wie nützlich, wie wichtig sind mir die Folgen einer Handlung?) und durch die Erwartungen, die man in einer Situation bezüglich der Folgen des eigenen Handelns hat (Heckhausen 1989). Heckhausen und Rheinberg (1980) haben hierzu ein Prozessmodell entwickelt, demzufolge sich eine Person in einer Lernsituation mehrere Fragen stellt, von deren Beantwortung es abhängt, ob eine Lerninitiative ergriffen wird.

Überträgt man die situationalen Parameter dieses Modells auf die Ebene stabiler Dispositionen, die sich als Ergebnis kumulativer Erfahrungen im Laufe der Schulzeit gebildet haben, dann ist das Pendant der Handlungs-Ergebnis-Erwartung die überdauernde Tendenz, sich Erfolge zuzutrauen und in Lernsituationen wirksam zu sein („Selbstwirksamkeit“). Theoretisch nahe damit verwandt ist das deklarative Wissen über die eigenen Kompetenzen und die eigene Leistungsfähigkeit, also das akademische Selbstkonzept.

Je nach Gegenstandsbereich lassen sich das akademische Selbstkonzept sowie die Lernmotivation aus verschiedenen Perspektiven modellieren (Helmke 1992):

- Fachübergreifend vs. schulfachspezifisch; letzteres ist inzwischen der Regelfall.
- Auf spezifische Kompetenzen bzw. Inhaltsbereiche (also eine Hierarchieebene unterhalb derjenigen des Gesamtfaches) vs. auf eine globale fachbezogene Fähigkeit oder einen Inhaltsbereich bezogen.
- Beim akademischen Selbstkonzept: Auf die aktuelle Situation vs. auf die Zukunft bezogen; in letzterem Falle geht es nicht um den Status, sondern um die Einschätzung des eigenen Potenzials.
- Alle angegebenen Aspekte lassen sich auf zwei unterschiedliche Vergleichsgruppen bzw. -normen beziehen – auf das absolute Niveau (also kriterial orientiert) oder auf den Vergleich mit Anderen, insbesondere mit der eigenen Schulklasse (also an sozialen Bezugsnormen orientiert).

Für das DESI-Projekt haben die Ergebnisse zum fachspezifischen Selbstkonzept sowie zur Lernmotivation eine mehrfache Relevanz:

- Ein positives akademisches Selbstkonzept und eine hohe Lernmotivation sind für sich genommen wichtige Bildungsziele (Weinert 1980), so dass beide Konzepte für DESI nicht nur als Bedingungsvariablen, sondern auch als Kriterien eine Rolle spielen.
- Eine der zentralen Fragestellungen von DESI ist die nach den individuellen Bedingungen der Sprachkompetenz im Deutschen und im Englischen. Aus der Pädagogischen Psychologie (Helmke/Weinert 1997) und zahlreichen Meta-Analysen wissen wir seit langem, dass das akademische Selbstkonzept und die Lernmotivation wichtige Bedingungen des Lernerfolges sind. Die Literatur hierzu ist nahezu unüberschaubar (vgl. die Übersicht bei Jerusalem 2006; Boekaerts 2002; Frenzel/Pekrun/Götz 2006; Helmke/Schrader 2006). Hier stellt sich somit die Frage, welchen Erklärungswert Unterschiede im Selbstkonzept bzw. der Lernmotivation für Unterschiede in den individuellen Testleistungen haben.

21.2 Instrumente

In Tabelle 21.1 sind diejenigen Skalen und Einzelitems aufgelistet, die in diesem Kapitel verwendet werden. Darüber hinausgehende Details sind dem Skalenhandbuch zu entnehmen¹.

Akademisches Selbstkonzept. Im Fach Deutsch wurde eine Skala zum leistungs- und fachbezogenen Selbstkonzept gebildet², die auf insgesamt vier Subskalen basiert: Je eine Skala zur Selbstwirksamkeit und zum fachbezogenen Selbstkonzept sowie zwei Skalen zum leistungsbezogenen Selbstkonzept. Beispielitems: „Ich bin überzeugt, dass ich das, was im Deutschunterricht vermittelt wird, beherrschen kann.“; „Wie schätzt du deine Beherrschung der deutschen Sprache ein? Richtige Grammatik beim Schreiben; die wichtigsten Inhalte eines Textes zusammenfassen“; „Egal wie sehr ich mich anstrengte, ich komme mit dem Fach Deutsch einfach nicht zurecht.“ (umgepolt); Antwortkategorien: *stimmt gar nicht / stimmt eher nicht / stimmt eher / stimmt ganz genau* (Selbstwirksamkeit, fachbezogenes Selbstkonzept); *sehr gut / gut / befriedigend / ausreichend / mangelhaft / ungenügend* (leistungsbezogenes Selbstkonzept).

Subjektive Testkompetenz Deutsch. Zusätzlich zu den auf einen größeren Zeitraum bezogenen (habituellen) leistungs- und fachbezogenen Selbstkonzepten wurde das aktuelle, auf die zum jeweiligen Messzeitpunkt in der neunten Jahrgangsstufe erhobenen Deutschtests bezogene Selbstkonzept erfasst. Beispielitem: „Du solltest Gespräche verfolgen und Argumente formulieren. Wie gut hast Du das gekonnt?“; Antwortkategorien: *sehr gut / gut / befriedigend / ausreichend / mangelhaft / ungenügend*.

¹ Alle Angaben hier sind – wie in Skalenhandbüchern üblich – ungewichtet berechnet worden.

² Cronbachs Alpha = .79 für die Gesamtskala zu Beginn bzw. .80 für die Skala am Ende der neunten Jahrgangsstufe.

Tabelle 21.1: Kennwerte der im Deutschen verwendeten Skalen und Items für die untersuchte Stichprobe (ungewichtete Angaben).

Konstrukt	Messzeitpunkt Jahrgangsstufe 9	M	SD	α	IKK ^A	Range	Stabi- lität
Personen- merkmale							
Akademisches Selbstkonzept ^B	Beginn	4.47	0.64	.79	.08	1-6	.66
	Ende	4.50	0.68	.80	.08		
Soziales Selbstkonzept	Beginn	4.60	0.75	.81	.02	1-6	.63
	Ende	4.61	0.80	.84	.02		
Lerninteresse	Beginn	2.32	0.57	.63	.08	1-4	.56
	Ende	2.27	0.62	.70	.08		
Leseinteresse	Beginn	2.77	0.92	.85	.10	1-4	.70
	Ende	2.66	0.97	.88	.10		
Literarische Orientierung	Ende	1.97	0.76	.83	.11	1-4	
Testbezogene Selbstein- schätzungen							
Subjektive Kompetenz	Beginn	3.80	0.85	.76	.11	1-6	.54
	Ende	3.95	0.84	.86	.14		
Valenz	Ende	3.01	0.83	.74	.08	1-5	
Testmotivation	Beginn	3.07	0.61	.84	.07	1-4	.41
	Ende	2.92	0.71	.90	.11		
Deutschnote	Ende ^C	2.89	0.85	-	.13	1-6	

Anmerkungen: ^AIKK = Intraklassenkorrelation; ^BGlobalwert aus den Subskalen Selbstwirksamkeit (reskaliert auf den Wertebereich 1-6), fachbezogenes Fähigkeitsselbstkonzept (reskaliert auf den Wertebereich 1-6) und zwei Skalen zum leistungsbezogenen Selbstkonzept; ^CErwartete Deutschnote am Ende der neunten Jahrgangsstufe.

Soziales Selbstkonzept. Neben den fachbezogenen Selbstkonzepten wurde das soziale Selbstkonzept erfasst. Beispielitem: „Wie gut kannst du in Diskussionen überzeugen?“; Antwortkategorien: sehr gut / gut / befriedigend / ausreichend / mangelhaft / ungenügend.

Im Bereich der Lernmotivation bzw. des -interesses wurden in DESI folgende Skalen zu Beginn und am Ende der neunten Jahrgangsstufe eingesetzt:

Lerninteresse für Deutsch. Die Skala umfasst die Einstellung zum Fach Deutsch sowie diesbezügliche freiwillige Lernaktivitäten. Beispielitem: „Ohne das Fach Deutsch würde mir etwas fehlen.“; Antwortkategorien: stimmt gar nicht / stimmt eher nicht / stimmt eher / stimmt ganz genau.

Leseinteresse. Hier geht es um das Lesen in der Freizeit und um die damit verbundenen Emotionen. Beispielitem: „Ich lese in meiner Freizeit.“; Antwortkategorien: stimmt gar nicht / stimmt eher nicht / stimmt eher / stimmt ganz genau.

Literarische Orientierung. Die Skala umfasst u.a. folgende Items: „Wie oft liest du folgende Texte zum Vergnügen? Abenteuerromane, Kriminalromane, historische Romane“.

Testmotivation. Nach der Bearbeitung sämtlicher DESI-Testmodule für den Bereich Deutsch zum jeweiligen Messzeitpunkt in der neunten Jahrgangsstufe wurde die Intensität der Anstrengung bei der Bearbeitung erfasst. Beispielitem: „Wie ist es dir bei den DESI-Tests ergangen? Ich habe mir Mühe gegeben.“; Antwortkategorien: *stimmt gar nicht / stimmt eher nicht / stimmt eher / stimmt ganz genau.*

Testbezogene Valenz. Unmittelbar nach der Bearbeitung der DESI-Testmodule wurden die Einstellungen der Schülerinnen und Schüler zu den jeweiligen Aufgaben erhoben. Für den Anfang der neunten Jahrgangsstufe konnte aufgrund mangelnder Reliabilität keine Skala gebildet werden. Beispielitem: „Du solltest verschiedene Texte lesen und Fragen dazu beantworten. Wie gerne hast Du diese Aufgabe bearbeitet?“; Antwortkategorien: *sehr gerne / ziemlich gerne / mittelmäßig gerne / nicht so gerne / ungern.*

21.3 Deskriptive Ergebnisse

Im folgenden Abschnitt werden zunächst Mittelwertunterschiede der einzelnen Selbstkonzept- und Motivationsskalen betrachtet, und zwar nach Bildungsgängen, Geschlecht und Erstsprache. Es folgen Darstellungen der Zusammenhänge dieser Skalen sowie ausgewählter Hintergrundmerkmale und Leistungen untereinander. Abschließend werden Ergebnisse zu den Lese- und Schreibaktivitäten der Schülerinnen und Schüler berichtet.

In Tabelle 21.2 sind Bildungsgangunterschiede der Selbstkonzept- und Motivationsskalen sowie der Zuwachs zwischen Anfang und Ende der neunten Jahrgangsstufe dargestellt. Mit Ausnahme der Testmotivation zu Beginn der neunten Jahrgangsstufe zeigen sich überall signifikante Unterschiede zwischen den Bildungsgängen, die teilweise auch von erheblicher praktischer Bedeutung sind.

Die höchsten Differenzen – etwa zwei Drittel der Standardabweichung (jeweils $p < .001$) – finden sich bei der subjektiven Testkompetenz sowie beim Leseinteresse zwischen den Bildungsgängen Hauptschule und Gymnasium. Beim Leseinteresse und der subjektiven Testkompetenz am Ende der neunten Jahrgangsstufe sind zudem alle Kontraste zwischen den einzelnen Bildungsgängen hoch signifikant³ – mit Ausnahme des Kontrastes zwischen den Bildungsgängen Realschule und Integrierte Gesamtschule. Die niedrigsten Ausprägungen finden sich jeweils im Bildungsgang Hauptschule, gefolgt von der Realschule sowie der Integrierten Gesamtschule, während im Gymnasium jeweils die höchsten Ausprägungen vorliegen.

3 $p < .01$ für den Kontrast bezüglich der subjektiven Deutschkompetenz in den Bildungsgängen Hauptschule und Integrierte Gesamtschule; $p < .001$ bei allen übrigen Kontrasten.

Tabelle 21.2: Selbstkonzept und Motivation in Bezug auf Deutsch nach Bildungsgang⁴ (z-standardisierte Werte⁵).

	Jahrgangsstufe 9	HS	RS	IGS	GY	ANOVA ^A	R ² (%)	Zuwachs (gesamt)
Akademisches Selbstkonzept	Anfang	-0.15	-0.06	-0.05	0.19	***	1.8	0.03 n.s.
	Ende	-0.19	-0.06	-0.04	0.33	***	3.8	
Soziales Selbstkonzept	Anfang	0.02	-0.05	0.00	0.04	*	0.2	0.00 n.s.
	Ende	0.00	-0.06	-0.05	0.08	***	0.3	
Lerninteresse	Anfang	0.16	-0.05	0.22	-0.09	***	1.2	-0.11***
	Ende	0.00	-0.19	0.09	-0.12	***	0.6	
Leseinteresse	Anfang	-0.31	-0.08	-0.05	0.33	***	6.1	-0.13***
	Ende	-0.47	-0.23	-0.10	0.24	***	6.6	
Testmotivation	Anfang	-0.04	-0.04	0.03	0.08	n.s.	0.3	-0.26***
	Ende	-0.54	-0.18	-0.30	-0.12	***	2.0	
Subjektive Testkompetenz Deutsch	Anfang	-0.24	-0.09	0.02	0.29	***	4.4	0.15***
	Ende	-0.18	0.09	0.10	0.50	***	6.7	
Testvalenz	Ende	-0.22	0.02	-0.02	0.15	***	2.0	

Anmerkungen: ^AANOVA mit Faktor Bildungsgang; Signifikanzniveaus: *** $p < .001$; ** $p < .01$; * $p < .05$.

Beim Lerninteresse zu Beginn der neunten Jahrgangsstufe hingegen ist diese Reihenfolge nahezu umgekehrt: Hier finden sich in der Hauptschule sowie in der Integrierten Gesamtschule die tendenziell positiveren Einstellungen, während das Fach Deutsch im Gymnasium und in der Realschule weniger geschätzt wird.

Bei der Testmotivation zeigt sich zu Beginn der neunten Jahrgangsstufe kein signifikanter Unterschied, während am Ende der neunten Jahrgangsstufe die Schülerinnen und Schüler im Bildungsgang Hauptschule deutlich weniger als in der Realschule bzw. im Gymnasium motiviert sind (jeweils $p < .001$). Der zuletzt genannte Sachverhalt lässt sich möglicherweise darauf zurückführen, dass für die meisten Hauptschüler nach der neunten Jahrgangsstufe die reguläre Schulzeit endet und für sie deshalb die Bearbeitung von Testaufgaben einen geringeren Anreizwert besitzt.

Betrachtet man die Veränderungen in den längsschnittlich erhobenen Skalen im Laufe der neunten Jahrgangsstufe, so zeigen sich hoch signifikante (jeweils $p < .001$) Abnahmen beim Lerninteresse, dem Leseinteresse (jeweils etwa eine zehntel Standardabweichung) sowie der Testmotivation⁶ (etwa eine viertel

4 In Tabellen und Abbildungen werden für die einzelnen Bildungsgänge folgende Abkürzungen verwendet: HS (Hauptschule), RS (Realschule), IGS (Integrierte Gesamtschule), GY (Gymnasium).

5 Bei den zu beiden Messzeitpunkten in der neunten Jahrgangsstufe erhobenen Konstrukten wurden jeweils der Mittelwert und die Streuung der Skala zu Beginn der neunten Jahrgangsstufe zur Standardisierung herangezogen.

6 Bei den auf die Tests bezogenen Skalen kann möglicherweise ein Teil der Veränderung auf die unterschiedlichen Testmodule zu beiden Messzeitpunkten in der neunten Jahrgangsstufe zurückgeführt werden.

Standardabweichung). Die testbezogene subjektive Deutschkompetenz steigt um knapp eine fünftel Standardabweichung ($p < .001$), während das akademische sowie das soziale Selbstkonzept im Niveau unverändert bleiben.

Das Bildungsgangprofil (aus Gründen der Übersichtlichkeit auf die Ergebnisse am Ende der neunten Jahrgangsstufe beschränkt) ist in Abbildung 21.1 dargestellt. Es handelt sich um die in Balkenform dargestellten z-transformierten mittleren Werte der jeweiligen Bildungsgänge. Die vom Nullpunkt nach links zeigenden Balken stehen für Werte unterhalb des Gesamtdurchschnitts, die nach rechts zeigenden Balken repräsentieren demzufolge überdurchschnittliche Ausprägungen.

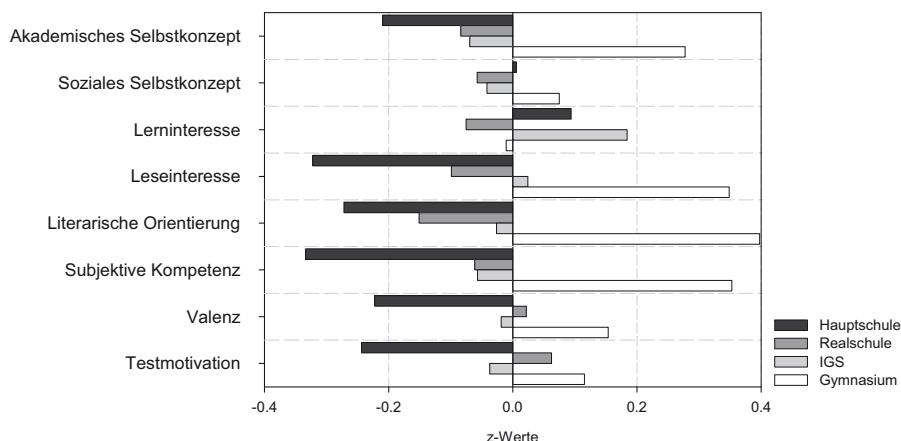


Abbildung 21.1: Bildungsgangspezifische Motivationsprofile in Bezug auf Deutsch.

Betrachtet man die Geschlechtsunterschiede, so zeigen sich bei allen hier untersuchten Konstrukten – Ausnahmen stellen lediglich das soziale Selbstkonzept sowie die subjektive Testkompetenz zu Beginn der neunten Jahrgangsstufe dar – günstigere Ausprägungen bei den Mädchen. Interessanterweise findet die – im Hinblick auf die höheren Testleistungen bei den Mädchen angemessene – höhere Ausprägung bezüglich des akademischen Selbstkonzeptes (etwa eine viertel Standardabweichung $p < .001$) bei den Mädchen zu Beginn der neunten Jahrgangsstufe keine Entsprechung im aktuellen testbezogenen Selbstkonzept. Am Ende der neunten Jahrgangsstufe ist dieser Unterschied dann zwar signifikant ($p < .05$), aber verglichen mit der noch extremeren Differenz von fast einer halben Standardabweichung hinsichtlich des habituellen Selbstkonzeptes doch eher zu vernachlässigen.

Die am stärksten ausgeprägten Geschlechtsunterschiede finden sich beim Leseinteresse: Dort liegen die Mädchen mehr als eine halbe Standardabweichung über den Jungen. Die bei PISA für den neunten Jahrgang gefundenen Geschlechtsunterschiede (Stanat/Kunter 2001) zeigen sich auch in DESI.

Betrachtet man die Unterschiede bezüglich der Erstsprache (Deutsch, nicht Deutsch, mehrsprachig), so liegen bei allen Konstrukten – mit Ausnahme des sozialen Selbstkonzeptes zum Beginn und dem Lerninteresse am Ende der neunten Jahrgangsstufe – signifikante Differenzen zwischen den drei Gruppen vor.

Vergleicht man das Motivationsprofil der deutschsprachigen mit den nicht-deutschsprachigen Schülerinnen und Schülern, dann zeigt sich, dass sich erstere in den Bereichen des Selbstkonzeptes im Deutschen, dem Leseinteresse und der Einstellung zu den Testaufgaben deutlich günstiger einschätzen – knapp eine fünftel bis fast eine halbe Standardabweichung – als ihre nicht-deutschsprachigen Mitschüler. Die mehrsprachigen Schülerinnen und Schüler hingegen unterscheiden sich kaum von den deutschsprachigen Schülerinnen und Schülern: Signifikante Unterschiede (jeweils $p < .05$) finden sich hier nur beim sozialen Selbstkonzept zum Ende der neunten Jahrgangsstufe, dem Lerninteresse am Anfang der neunten Jahrgangsstufe – dort liegen die mehrsprachigen Schülerinnen und Schüler knapp über den deutschsprachigen – sowie – in umgekehrter Reihenfolge – der Testmotivation und der subjektiven Testkompetenz zum Ende der neunten Jahrgangsstufe.

In Tabelle 21.3 sind die Interkorrelationen der Selbstkonzept- und Motivationskalen, der literarischen Orientierung, der Deutschleistung bzw. -note sowie zentraler Hintergrundmerkmale getrennt nach Individual- (also innerhalb von Schulklassen) und Klassenebene (d.h. zwischen Klassen) dargestellt. Die Zusammenhänge der Selbstkonzept- und Motivationsskalen liegen auf der Individualebene zwischen $r = .04$ (soziales Selbstkonzept und Leseinteresse) und $r = .60$ (subjektive Kompetenz und Testvalenz), auf der Klassenebene zwischen $r = -.13$ (Testmotivation und Lerninteresse) und $r = .83$ (Subjektive Kompetenz und Leseinteresse). Die Skala „Leseinteresse“ korreliert – auf beiden Ebenen – sehr hoch mit der literarischen Orientierung. Die höchsten Zusammenhänge sowohl mit der Deutschnote als auch mit der Deutschleistung auf Individualebene weist das akademische Selbstkonzept auf. Auch das Leseinteresse hängt stark mit der Deutschleistung zusammen, insbesondere auf der Klassenebene ($r = .91$). Hier zeigen sich auffällig hohe Interkorrelationen zwischen dem Leseinteresse, der literarischen Orientierung, der Deutschleistung, dem Sozialstatus und der kognitiven Grundfähigkeit, die im Bereich von $r = .78$ bis $r = .95$ liegen und zum Teil Bildungsgangunterschiede repräsentieren. Ausgeprägte Unterschiede zwischen Zusammenhängen auf Klassen- und Individualebene zeigen sich vor allem beim Leseinteresse bzw. der literarischen Orientierung und dem Sozialstatus: Die Zusammenhänge zwischen Schulklassen liegen bei $r = .94$ (Leseinteresse) bzw. $r = .95$ (literarische Orientierung), während sie innerhalb von Klassen bei $r = .06$ bzw. $r = .07$ liegen.

Tabelle 21.3: Interkorrelationen^A der Skalen^B und Zusammenhänge mit Hintergrundvariablen bzw. Deutschleistung getrennt nach Individualebene (oberhalb der Diagonale) und Klassenebene (unterhalb der Diagonale) in Bezug auf Deutsch.

	Akademisches SK	Soziales SK	Lerninteresse	Leseinteresse	Literarische Orientierung	Subjektive Kompetenz	Testvalenz	Testmotivation	Deutschnote	Deutschleistung	Sozialstatus	kognitive Grundfähigkeit	Geschlecht	Erstsprache Deutsch
Akademisches SK		.43	.38	.31	.24	.43	.29	.20	.49	.42	.05	.09	.22	.10
Soziales SK	.51		.10	.04	.05	.22	.10	.08	.18	.07	.03	-.03		
Lerninteresse	.30			.24	.25	.18	.31	.18	.23	.06		-.06	.13	.02
Leseinteresse	.70	.42			.70	.19	.22	.17	.19	.34	.06	.09	.29	.04
Literarische Orientierung	.72	.57		.95		.15	.16	.12	.16	.22	.07	.08	.22	
Subjektive Kompetenz	.76	.64		.83	.83		.60	.26	.22	.27	.05	.05	.03	.12
Testvalenz	.49	.63		.61	.52	.77		.29	.14	.19		.04	.13	.07
Testmotivation	.35			.44	.32	.54	.79		.08	.21		.08	.13	.04
Deutschnote ^C	.58		.22	.31	.32	.24	.24			.34	.05	.11	.17	.06
Deutschleistung	.68	.42	-.21	.91	.84	.77	.55	.55	.24		.08	.33	.21	.24
Sozialstatus	.72	.40		.94	.89	.79	.48	.39	.35	.93		.03	-.06	.14
kognitive Grundfähigkeit	.61	.32	-.19	.83	.78	.72	.53	.47	.29	.93	.87			.05
Geschlecht ^D	.20		-.22	.33	.18	.13	.17	.23		.34	.22	.22		
Erstsprache Deutsch ^E			-.29	.50	.28	.28	.24	.21	.23	.50	.52	.52	.15	

Anmerkungen: ^ABasis: Zwei-Ebenen-Analysen mit Mplus 4.0; nicht signifikante Korrelationen ($p \geq .05$; einseitige Tests) sind nicht dargestellt; ^BAlle hier dargestellten Skalen wurden am Ende der neunten Jahrgangsstufe erhoben; ^CErwartete Deutschnote am Ende der neunten Jahrgangsstufe (umgepolt). Es sei darauf hingewiesen, dass die Deutschnote auf Klassenebene aufgrund der unterschiedlichen Benotungsmaßstäbe in den verschiedenen Bildungsgängen hier kaum interpretierbar und nur der Vollständigkeit halber dargestellt ist; ^DKodierung: 0 = Junge, 1 = Mädchen; ^EKodierung: 0 = nicht Deutsch, 1 = Deutsch.

Lese- und Schreibaktivitäten

Bei DESI wurde auch danach gefragt, was und wie oft außerhalb der Schule („zum Vergnügen“) in deutscher Sprache gelesen bzw. geschrieben wird. Da bisher – außer in den nationalen Jugendstudien – wenig Informationen hierzu vorliegen, sollen einige der Ergebnisse berichtet werden (vgl. dazu Abbildungen 21.2 und 21.3). Dies er-

scheint auch deshalb wichtig, weil sie Hinweise auf die Verknüpfung von Unterricht und Lebenswelt der Schülerinnen und Schüler liefern können.

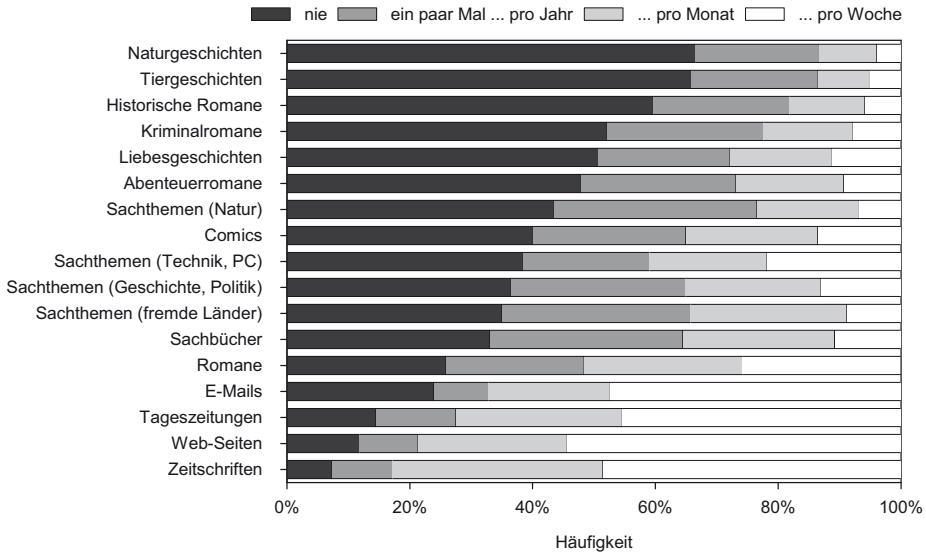


Abbildung 21.2: Prozentuale Anteile der Häufigkeiten des Lesens verschiedener Textsorten in deutscher Sprache in der Freizeit.

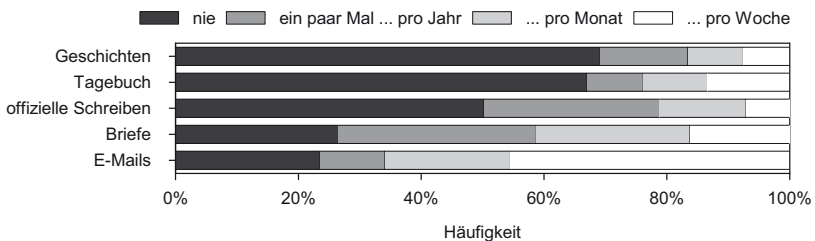


Abbildung 21.3: Prozentuale Anteile der Häufigkeiten der Produktion verschiedener Textsorten in deutscher Sprache in der Freizeit.

Der größten Beliebtheit erfreuen sich Zeitschriften, Internet-Seiten, Tageszeitungen und E-Mails. Für jeweils etwa die Hälfte der Schülerinnen und Schüler gehört das Lesen solcher Texte zum Alltag. Betrachtet man das Lesen von Romanen, so verteilen sich die Angaben der Schülerinnen und Schüler in etwa zu gleich großen Teilen auf die vier Antwortkategorien. Im mittleren Bereich finden sich die Sachbücher: Etwa zwei Drittel aller Schülerinnen und Schüler lesen zumindest gelegentlich Sachbücher in ihrer Freizeit. Am seltensten von den hier erfassten Textsorten werden Tier- bzw. Naturgeschichten gelesen.

Die häufigste der in DESI erhobenen schriftlichen Formen der Textproduktion ist das Schreiben von E-Mails: Knapp die Hälfte der Schülerinnen und Schüler geben an, mehrmals pro Woche E-Mails zu verfassen. Relativ häufig – mindestens ein paar Mal pro Monat bei 41% der Schülerinnen und Schüler – werden auch Briefe geschrieben. Das Führen eines Tagebuchs und das Verfassen eigener Geschichten sind eher wenig

verbreitete Formen der schriftlichen Textproduktion, die nur bei etwa einem Drittel der Schülerinnen und Schüler wenigstens gelegentlich vorkommen. Erstaunlich ist allerdings, dass etwa 8% der Schülerinnen und Schüler eigenen Angaben zufolge mehrmals pro Woche „Geschichten für sich selbst“ schreiben.

21.4 Motivation und Leistungen

Tabelle 21.4: Partialkorrelationen⁷ von Selbstkonzept- und Motivationsskalen^A mit der Deutschleistung (am Ende der neunten Jahrgangsstufe).

	Deutsch-Testleistung			Deutsch-Zeugnisnote ^B		
	Individual- ebene	innerhalb von Klassen	zwischen Klassen	Individual- ebene	innerhalb von Klassen	zwischen Klassen
Akademisches Selbstkonzept	.19***	.31***	.00	.33***	.33***	.33***
Soziales Selbstkonzept	.03**	.03**	-.01	.16***	.15***	.31**
Lerninteresse	.00	.03	-.03	.14***	.12***	.24**
Leseinteresse	.15***	.23***	.02	.11***	.12***	.01
Testmotivation	.10***	.14***	.10**	.02	.05*	-.12
Subjektive Testkompetenz Deutsch	.13***	.19***	.01	.16***	.18***	-.02
Einstellung zu Testaufgaben	.10***	.13***	.04	.11***	.10***	.13*

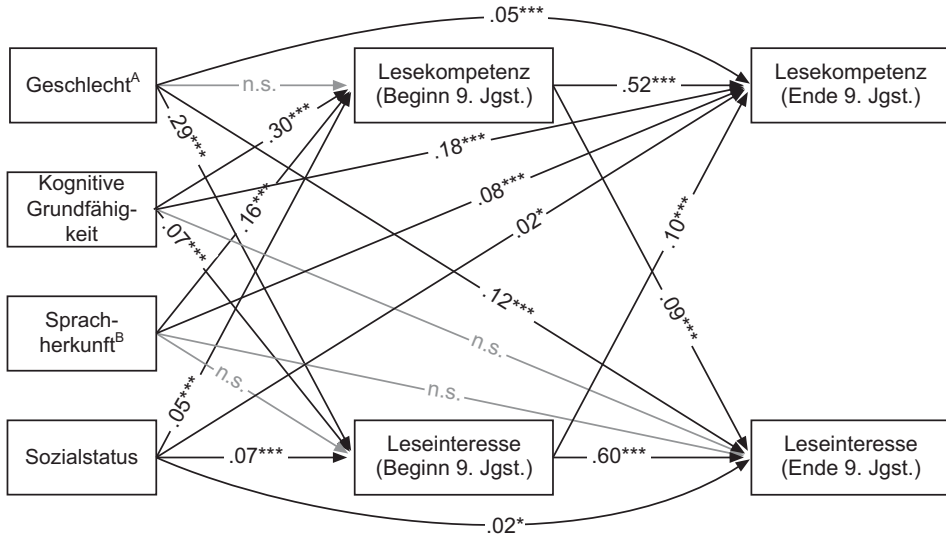
Anmerkungen: ^AJeweils zu Beginn der neunten Jahrgangsstufe erhoben (Ausnahme: Einstellung zu den Testaufgaben); ^BErwartete Zeugnisnote am Ende der neunten Jahrgangsstufe, umgepolt; Signifikanzniveaus: *** ($p < .001$); ** ($p < .01$); * ($p < .05$).

Betrachtet man die um die Eingangsvoraussetzungen bereinigten Zusammenhänge zwischen den hier betrachteten Selbstkonzept- bzw. Motivationsskalen und der Deutschleistung im einfachen Regressionsmodell bzw. auf der Individualebene im Zwei-Ebenen-Modell (vgl. Tabelle 21.4), so zeigen sich im Wesentlichen die erwarteten positiven Korrelationen. D.h. Schülerinnen und Schüler, die zu Beginn der neunten Jahrgangsstufe (unter Berücksichtigung zentraler Ausgangsbedingungen) über ein hohes – vor allem fachliches, aber auch soziales – Selbstkonzept verfügen bzw. hoch motiviert sind, erbringen am Ende der neunten Jahrgangsstufe höhere

7 Berichtet werden Partialkorrelationen auf der Basis eines Standard-Modells (Individualebene), bei dem die Mehrebenenstruktur der Daten lediglich bei der Schätzung der Standardfehler berücksichtigt wird, sowie eines Zwei-Ebenen-Modells, bei dem Partialkorrelationen innerhalb von Schulklassen (Ebene 1) und zwischen Klassen (Ebene 2) simultan geschätzt werden. Alle Zusammenhangsanalysen in diesem Kapitel wurden mit Mplus 4.0 gerechnet. Kontrolliert wurden jeweils folgende Hintergrundmerkmale: Sozialstatus, kognitive Grundfähigkeit, Geschlecht und Erstsprache (Deutsch vs. andere Sprache); im Standard-Modell und auf Klassenebene zusätzlich Bildungsgang.

Leistungen. Eine Ausnahme stellt lediglich das Lerninteresse dar, das nicht mit der Testleistung, sondern nur mit der Deutschnote zusammenhängt.

Auf der (Schul-) Klassenebene ist bezüglich der Testleistung lediglich die Testmotivation von Bedeutung: Eine hohe Testmotivation geht mit hoher Testleistung gegen Ende der neunten Jahrgangsstufe einher. Nimmt man die durchschnittliche erwartete Deutschnote am Ende der neunten Jahrgangsstufe als Kriterium, dann sind vor allem das Niveau des leistungs- und fachbezogenen sowie des sozialen Selbstkonzepts, des Lerninteresses bzw. der Einstellung zu den Testaufgaben von Bedeutung⁸.



Anmerkung: ^A0 = Junge, 1 = Mädchen; ^B0 = nicht Deutsch, 1 = Deutsch.

Abbildung 21.4: Pfadmodell über den Zusammenhang zwischen Hintergrundvariablen, Leseinteresse und Lesekompetenz zu Beginn und am Ende der neunten Jahrgangsstufe (innerhalb von Klassen) in Bezug auf Deutsch⁹.

Die Stabilitäten und die wechselseitige Beeinflussung der Lesekompetenz und des Leseinteresses (innerhalb von Schulklassen) unter Kontrolle zentraler Hintergrundmerkmale sind in Abbildung 21.4 dargestellt. Es zeigen sich erwartungsgemäß relativ hohe Stabilitäten bei beiden Konstrukten, die durch den Einbezug der Hintergrundvariablen etwas abgeschwächt werden. Entscheidend sind insbesondere die beiden sogenannten *cross-lagged*-Pfade zwischen der Lesekompetenz zu Beginn der neunten Jahrgangsstufe und dem Leseinteresse zum Ende der neunten Jahrgangsstufe bzw. dem Leseinteresse zu Beginn und der Lesekompetenz

8 Möglicherweise spielt hier die „Erwartung“ bezüglich der Zeugnisnote auch eine entscheidende Rolle: Es wäre denkbar, dass Schülerinnen und Schüler bzw. Klassen mit hohem (akademischen) Selbstkonzept oder hohem Lerninteresse zu einer zu optimistischen Notenerwartung – also zu einer Überschätzung – neigen.

9 Im Modell wurden alle Zusammenhänge spezifiziert. Aus Gründen der Übersichtlichkeit sind sie jedoch nicht alle in der Abbildung dargestellt.

am Ende der neunten Jahrgangsstufe. Beide Pfade sind hoch signifikant. Dieses Ergebnis kann folgendermaßen interpretiert werden¹⁰: Ein – jeweils bezogen auf den Klassendurchschnitt unter Kontrolle der vier Hintergrundmerkmale relativ – hohes Leseinteresse führt zu einem Zuwachs¹¹ der Lesekompetenz. Und auch umgekehrt führt eine hohe Lesekompetenz zu einem Zuwachs im Leseinteresse. Dies könnte im Laufe der Zeit zu einem so genannten „Schereneffekt“ führen: Der Abstand bezüglich der Lesekompetenz zwischen Schülerinnen und Schülern könnte sich zunehmend – vermittelt über das Leseinteresse – vergrößern. Durch eine Förderung des Leseinteresses (insbesondere bei eher leistungsschwachen Schülerinnen und Schülern) ließe sich aber möglicherweise einem solchen „Schereneffekt“ entgegenwirken.

Betrachtet man die von den Hintergrundmerkmalen ausgehenden Pfade, so sind insbesondere folgende Zusammenhänge erwähnenswert: Mädchen zeigen ein deutlich höheres Leseinteresse als Jungen bei gleichzeitig höherem Zuwachs. Die kognitive Grundfähigkeit spielt vor allem bei der Lesekompetenz eine Rolle. Hohe Ausprägungen führen¹² sowohl zu hoher Kompetenz als auch zu einem höheren Zuwachs. Gleiches gilt für die Sprachherkunft (Erstsprache Deutsch), wenngleich in geringerem Umfang. Bezogen auf das Leseinteresse ist die Erstsprache erstaunlicherweise ohne Bedeutung. Der Sozialstatus ist sowohl mit dem Status als auch dem Zuwachs beider Konstrukte signifikant positiv – wenngleich bezüglich der Größenordnung der standardisierten Pfadkoeffizienten ($\beta = .02$ bis $\beta = .07$) nur sehr schwach – verknüpft.

Erweitert man dieses Modell um den häuslichen Buchbesitz (hier nicht dargestellt), der in DESI zu beiden Messzeitpunkten in der neunten Jahrgangsstufe erhoben wurde, so zeigen sich erwartungsgemäß positive Zusammenhänge zwischen dem Buchbesitz zu Beginn der neunten Jahrgangsstufe und dem Zuwachs im Leseverstehen bzw. -interesse ($\beta = .04$ bzw. $\beta = .03$). Bemerkenswert sind aber insbesondere die positiven Pfade von Lesekompetenz bzw. -interesse auf den Zuwachs bezüglich des Buchbesitzes ($\beta = .04$ bzw. $\beta = .09$). Das heißt: Der häusliche Buchbesitz ist keine einfache „Hintergrundvariable“¹³ im Sinne einer von Schülerinnen und Schülern unabhängigen Lernvoraussetzung. Schülerinnen und Schüler nehmen offensichtlich abhängig von ihrer Lesekompetenz und ihrem Leseinteresse Einfluss auf den Buchbesitz zu Hause, möglicherweise durch die Anregung, bestimmte Bücher anzuschaffen, oder durch den Kauf eigener Bücher.

10 Eine zwingende kausale Verknüpfung lässt sich daraus nicht ableiten.

11 Da es sich hier um korrelative Zusammenhänge handelt, muss „Zuwachs“ im Sinne der relativen Position innerhalb der Klasse – und nicht zwingend im tatsächlichen Niveau – interpretiert werden.

12 Da die kognitive Grundfähigkeit am Ende der neunten Jahrgangsstufe erhoben wurde, setzt diese Interpretation allerdings eine relativ starke Stabilitätsannahme voraus.

13 Als solche wird der Buchbesitz häufig im Sinne eines Indikators der sozialen Herkunft verwendet.

21.5 Ausblick

In diesem Kapitel konnten lediglich die grundlegenden Ergebnisse aus den Bereichen Selbstkonzept und Motivation im Fach Deutsch dargestellt werden. Die umfangreiche Datenbasis des Projektes DESI erlaubt darüber hinaus eine Reihe von vertiefenden Analysen, auf die aus Platzgründen hier nicht weiter eingegangen werden konnte. Es sei an dieser Stelle auf den Abschnitt „Ausblick“ im Kapitel 22 verwiesen, wo die analogen Analysen für das Fach Englisch dargestellt sind.

Literatur

- Boekaerts, M. (2002): Motivation to learn (Educational Practices Series, 10). <http://www.ibe.unesco.org>.
- Frenzel, A. C./Pekrun, R./Götz, T. (2006): Emotionale Voraussetzungen des Lernens. In: Arnold, K.-H./Sandfuchs, U./Wiechmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt, S. 579-583.
- Heckhausen, H. (1989): Motivation und Handeln. Heidelberg: Springer.
- Heckhausen, H./Rheinberg, F. (1980): Lernmotivation im Unterricht, erneut betrachtet. In: Unterrichtswissenschaft 8, S. 7-47.
- Helmke, A. (1992): Selbstvertrauen und schulische Leistungen. Göttingen: Hogrefe.
- Helmke, A./Schrader, F.-W. (2006): Determinanten der Schulleistung. In: Rost, D. H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: Psychologie Verlags Union, S. 83-94.
- Helmke, A./Weinert, F. E. (1997): Bedingungsfaktoren schulischer Leistungen. In: Weinert, F. E. (Hrsg.): Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Pädagogische Psychologie, Vol. 3). Göttingen: Hogrefe, S. 71-176.
- Jerusalem, M. (2006): Motivationale und volitionale Voraussetzungen des Unterrichts. In: Arnold, K.-H./Sandfuchs, U./Wiechmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt, S. 575-579.
- Stanat, P./Kunter, M. (2001): Geschlechterunterschiede in Basiskompetenzen. In: Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K. J./Weiß, M. (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich, S. 251-269.
- Weinert, F. E. (1980): Lernmotivation – Psychologische Forschung und pädagogische Aufgabe. In: Unterrichtswissenschaft 8, S. 197-205.

Andreas Helmke / Friedrich-Wilhelm Schrader / Wolfgang Wagner /
Günter Nold / Konrad Schröder

22 Selbstkonzept, Motivation und Englischleistung

In Kapitel 21 (Wagner u.a.) ist bereits der theoretische Hintergrund zur Rolle motivationaler Merkmale für schulische Leistungen im Fach Deutsch skizziert worden. Dort wurde auch darauf hingewiesen, dass motivationale Orientierungen und sprachbezogene Lernaktivitäten zugleich Bedingungen und Folgen des schulischen Unterrichts sind (Weinert 1980). Auch für dieses Kapitel gilt, dass das fachspezifische Selbstkonzept und das fachspezifische Interesse - also Erscheinungsformen der Erwartung und des Wertes - die entscheidenden motivationalen Parameter sind (Corbière u.a. 2006). Im Folgenden wird deshalb nur auf Besonderheiten hingewiesen, die für das Fach Englisch gelten.

Anders als für den Deutschunterricht (Bremerich-Vos/Schnaitmann 1995) gibt es für die Motivation im Fach Englisch eine gut etablierte Tradition empirischer Forschung, allerdings überwiegend im angloamerikanischen Sprachraum (Dörnyei 2000; Dörnyei/Schmidt 2001). Aber auch im deutschen Sprachraum zeigt sich ein Bild lebendiger Forschung (Finkbeiner 2003; Nold/Schnaitmann 1995; Grotjahn 2003; Solmecke/Boosch 1981; Solmecke 1983; Nold 2003; Nold/Haudeck/Schnaitmann 1997) und eine rege Diskussion über fachdidaktische Fragen (Butzkamm 2004). Allerdings ist die Forschung im deutschsprachigen Raum eher qualitativ ausgerichtet, wie z.B. das Themenheft „Motivation“ der *Zeitschrift für Interkulturellen Fremdsprachenunterricht* (Riemer/Schlag 2004) belegt. Motivationale Merkmale im Fach Englisch wurden auch in Large-Scale-Studien wie in der IEA-Studie „Teaching English as a Foreign Language“ (vgl. hierzu Helmke/Schrader 2001), durch Wienold u.a. (1985), in den Konstanzer Studien (Fend 1982), in den Projekten TOSCA (Köller/Trautwein 2004) und BIJU (Gruehn 2000) untersucht. Die Datenlage bei DESI unterscheidet sich von den meisten der genannten Studien dadurch, dass es sich um eine repräsentative und längsschnittlich angelegte Untersuchung handelt.

22.1 Deskriptive Ergebnisse

Beschreibung der Instrumente

Die für dieses Kapitel relevanten Instrumente sind in Tabelle 22.1 aufgelistet. Zu jedem Konstrukt werden folgende Kennwerte berichtet: Messzeitpunkt, Mittelwert (M), Standardabweichung (SD), Reliabilität (Cronbachs α), Intraklassen-Korrelation (als Indikator des Ausmaßes von Schulklassenunterschieden), Wertebereich (Range)

und Stabilität (sofern das Merkmal zu beiden Messzeitpunkten in der neunten Jahrgangsstufe erhoben wurde).

Tabelle 22.1: Kennwerte der auf das Fach Englisch bezogenen Skalen und Items^A.

Konstrukt	Messzeitpunkt 9. Jgst.	M	SD	α	IKK	Range	Stabilität
Personmerkmale							
Akademisches Selbstkonzept ^B	Beginn	4.32	0.81	.84	.11	1-6	.76
	Ende	4.30	0.88	.84	.11		
Lerninteresse	Beginn	2.57	0.65	.79	.13	1-4	.67
	Ende	2.49	0.69	.81	.12		
Testbezogene Selbst-einschätzungen							
Subjektive Kompetenz	Beginn	3.52	0.79	.81	.26	1-6	.62
	Ende	3.51	0.86	.85	.25		
Valenz	Ende	2.64	0.77	.65	.13	1-5	-
Testmotivation	Beginn	3.08	0.61	.84	.09	1-4	.42
	Ende	2.95	0.71	.91	.10		
Außerschulische sprachrelevante Lernaktivitäten							
Engl. TV, Video	Beginn	1.83	0.74	.75	.07	1-4	.58
	Ende	1.84	0.77	.78	.08		
Englische Texte	Beginn	1.93	0.58	.79	.13	1-4	.58
	Ende	1.96	0.62	.82	.08		
Englischnote	Beginn ^C	3.08	0.95	-	.16	1-6	.54
	Ende ^D	3.00	0.92	-	.12		
Hausaufgabenzeit für Englisch (werktags)	Beginn	3.68	1.64	-	.06	1-8	.25
	Ende	3.17	1.66	-	.05		

Anmerkungen: ^AAlle Angaben beziehen sich auf die Stichprobe, nicht auf die Population. Wenn in der Spalte „ α “ keine Angaben sind, handelt es sich um Einzelitems;

^BGlobalwert aus den Subskalen Selbstwirksamkeit (reskaliert auf den Wertebereich 1-6), fachbezogenes Fähigkeitsselbstkonzept und leistungsbezogenes Selbstkonzept (reskaliert auf den Wertebereich 1-6); ^CNote am Ende der achten Jahrgangsstufe;

^DErwartete Note am Ende der neunten Jahrgangsstufe.

Das *akademische Selbstkonzept* ist eine aus drei reliablen Skalen gebildete Gesamtskala. Sie umfasst die eigene *Fähigkeit* für das Fach Englisch („Für Englisch habe ich einfach keine Begabung“), *Selbstwirksamkeit* („Wenn ich mir in Englisch Mühe gebe, dann kann ich es auch“) und die Einschätzung *spezifischer fachlicher Leistungen und Kompetenzen* („Wie gut bist du in der Lage, flüssig und ohne zu stocken Englisch zu sprechen?“). Die Skala zum *Lerninteresse* vereinigt Fragen zur fachbezogenen Einstellung („Ohne Englisch würde mir etwas fehlen“) und zu frei-

willigen *Lernaktivitäten* im Fach Englisch („Ich mache für Englisch mehr, als ich für die Schule brauchen würde“).

Neben diesen überdauernden Personmerkmalen („traits“) wurden unmittelbar nach der Durchführung jedes Subtests Fragen gestellt, die sich auf die aktuell vorliegende Ausprägung („state“) der subjektiven Kompetenz und der Valenz der bearbeiteten Aufgaben bezogen. Die Skala *testbezogene subjektive Kompetenz* enthält Items wie: „Du hast einen englischsprachigen Text gehört, zu dem Du Fragen beantworten solltest – Wie gut hast du das gekonnt?“ (Antwortkategorien: sehr gut / gut / eher gut / eher schlecht / schlecht / sehr schlecht). Bei der Skala *testbezogene Valenz* wurde gefragt, wie gerne die entsprechenden Aufgaben bearbeitet worden seien (Antwortkategorien: sehr gerne / ziemlich gerne / mittelmäßig gerne / nicht so gerne / ungern). Bei der *Testmotivation* geht es um das berichtete Ausmaß der Anstrengung und Konzentration während der Leistungstests.

Um etwas über *außerschulische*, für Englisch potenziell relevante Aktivitäten zu erfahren, wurden die Schüler danach gefragt, ob und wie häufig sie sich zum Vergnügen (also über schulische Anforderungen hinaus) mit englischsprachigen Medien (wie TV, Video, Email) und Texten (wie Romane, Manuale, Songs, Sachbücher und Comics) beschäftigen. Zu beiden Aspekten wurden Skalen gebildet.

Als *Englishtestleistung* wurde ein Globalwert verwendet, in den alle Teilkompetenzen mit gleichem Gewicht eingehen.

Neben diesen Skalen wurden mit Einzelitems erfasst: die Englisch-Zeugnisnoten (Ende achte Jahrgangsstufe, Zwischenzeugnis Mitte der neunten Jahrgangsstufe und erwartete Note für das Ende der neunten Jahrgangsstufe) sowie die Dauer der täglichen Hausaufgabenzeit für Englisch.

Tabelle 22.2¹ enthält die Interkorrelationen der Skalen und Items, die oben genannt wurden und in diesem Kapitel berücksichtigt werden. Oberhalb der Diagonale sind die Korrelationen auf der Individualebene innerhalb von Klassen (d.h. Klassenunterschiede im Niveau werden hier ausgeklammert); unterhalb der Diagonale sind die Korrelationen auf der Schulklassenebene aufgelistet.

1 Selbstkonzept, Motivation Englisch: Interkorrelationsmatrix (2-Ebenen); *Innerhalb* von Klassen (within): oberhalb der Diagonale; *Zwischen* Klassen (between): unterhalb der Diagonale; Basis: bivariate Korrelationen mit Mplus (Twolevel). Alle Variablen wurden am Ende der neunten Jahrgangsstufe erhoben. Signifikanzangaben: einseitige Tests.

Tabelle 22.2: Interkorrelation der Variablen auf Klassenebene (unterhalb der Diagonalen) und Individualebene (oberhalb der Diagonalen). Nicht signifikante Korrelationen ($p > .05$) sind nicht dargestellt.

innerhalb von Klassen zwischen Klassen	Akademisches SK	Lerninteresse	Subjektive Kompetenz	Valenz	Testmotivation	TV, Video	Texte	Hausaufgabenzeit	Erwartete Englischnote ^B	Englischleistung	Sozialstatus	Kognitive Grundfähigkeit	Geschlecht ^C	Erstsprache ^D
Akademisches SK		.61	.52	.37	.13	.26	.26		.58	.56	.04		.04	
Lerninteresse	.86		.35	.36	.18	.38	.39	.10	.38	.38	.04		.07	
Subj. Kompetenz	.73	.67		.54	.20	.17	.20		.33	.37		.06	-.08	-.04
Valenz	.58	.65	.78		.28	.18	.19	.06	.23	.27		.06	.09	-.05
Testmotivation	.41	.37	.48	.68		.04	.07	.16		.12		.07	.18	.07
TV, Video	.56	.68	.73	.57	.27		- ^A	.08	.13	.14	.07		-.07	
Texte	.60	.70	.76	.63	.34	- ^A		.14	.14	.15	.07			
Hausaufgabenzeit			.27	.26					-.04	-.05			.11	
Erw. Englischnote ^B	.49	.40		.16				-.40		.45	.05	.09	.08	
Englischleistung	.69	.62	.79	.68	.52	.68	.69		.15		.05	.26	.14	
Sozialstatus	.78	.60	.66	.54	.42	.63	.67		.26	.92			-.06	.15
Kogn. Grundfähigk.	.65	.51	.60	.57	.50	.58	.64		.20	.88	.86			.05
Geschlecht ^C	.25	.20		.19	.26			.24		.31	.23	.26		
Erstsprache ^D	.23							-.21	.19	.38	.56	.52	.19	

Anmerkungen: ^ADas entsprechende Modell konvergierte nicht; ^Bumgepolt; ^C0 = Junge, 1 = Mädchen; ^D0 = nicht Deutsch, 1 = Deutsch.

Exemplarisch für die Unterschiede zwischen der Individual- und der Klassenebene sei lediglich auf zwei Zusammenhänge hingewiesen: Auf der Schulklassenebene korreliert die Testmotivation mit der Englischtestleistung hoch (.52), auf der Individualebene dagegen deutlich geringer (.12). Mit anderen Worten: Für die Englischleistung macht es einen Unterschied, ob die Klasse insgesamt mehr oder weniger motiviert ist, weniger dagegen, ob sich Schülerinnen und Schüler innerhalb einer Klasse hinsichtlich ihrer Testmotivation unterscheiden. Beim Zusammenhang zwischen der Englischnote und der Englisch-Testleistung ist das Verhältnis umgekehrt: Interindividuelle Unterschiede in der Englischnote innerhalb einer Klasse spiegeln Unterschiede in der Testleistung wider (.45), während die durchschnittliche Englischnote von Klassen zwar ebenfalls signifikant (.15), aber erwartungsgemäß deutlich schwächer mit der durchschnittlichen Testleistung zusammenhängt. Vor einer Überinterpretation dieser Korrelationen wird allerdings gewarnt, weil es sich

um bivariate Zusammenhänge handelt; Hintergrundmerkmale (wie Bildungsgang, Geschlecht, kognitive Grundfähigkeit, Erstsprache) sind nicht herausgerechnet.

Motivationsprofile nach Bildungsgang, Geschlecht und Erstsprache

Tabelle 22.3 berichtet Bildungsgangunterschiede der Selbstkonzept- und Motivationsskalen sowie den Zuwachs zwischen Beginn und Ende der neunten Jahrgangsstufe. Wie im Fach Deutsch zeigt sich auch im Fach Englisch eine signifikante Abnahme beim Lerninteresse und - vor allem im Bildungsgang Hauptschule - noch stärker bei der Testmotivation, dagegen eine leichte Steigerung des Interesses an englischsprachigen Texten.

Tabelle 22.3: Selbstkonzept und Motivation Englisch nach Bildungsgang (z-standardisierte Werte).

	Jahrgangsstufe 9	HS	RS	IGS	GY	ANOVA ^A	R ² (%)	Zuwachs (gesamt)
Akademisches Selbstkonzept	Anfang	-0.11	-0.14	-0.39	0.29	***	4.5	
	Ende	-0.12	-0.18	-0.35	0.32	***	4.8	n.s.
Lerninteresse	Anfang	-0.11	-0.12	-0.24	0.26	***	3.3	
	Ende	-0.25	-0.23	-0.28	0.17	***	3.2	-.11***
Subjektive Kompetenz	Anfang	-0.30	-0.08	-0.34	0.39	***	7.8	
	Ende	-0.31	-0.12	-0.22	0.40	***	6.8	n.s.
Valenz	Ende	-0.22	-0.02	-0.17	0.22	***	2.8	
Englischsprachige Medien	Anfang	-0.14	-0.05	-0.07	0.17	***	1.6	
	Ende	-0.20	-0.07	0.06	0.24	***	2.7	n.s.
Englische Texte	Anfang	-0.24	-0.08	-0.19	0.30	***	4.7	
	Ende	-0.15	-0.01	0.09	0.33	***	2.8	0.07***
Testmotivation	Anfang	-0.17	0.00	-0.06	0.13	***	1.3	
	Ende	-0.54	-0.14	-0.31	-0.07	***	2.5	-0.22***

Anmerkungen: ^AANOVA mit Faktor Bildungsgang (WesVar, Fay-Replicates); *** $p < .001$.

Abbildung 22.1 gibt das Bildungsgangprofil der ausgewählten motivationalen Skalen wieder (beschränkt auf die Ergebnisse am Ende der neunten Jahrgangsstufe). Da sich die Metrik der Skalen unterscheidet, werden in der Abbildung z-standardisierte Werte verwendet. Der Durchschnitt (über alle Bildungsgänge hinweg) liegt immer bei Null; Balken im positiven Zahlenbereich verkörpern also überdurchschnittliche Werte.

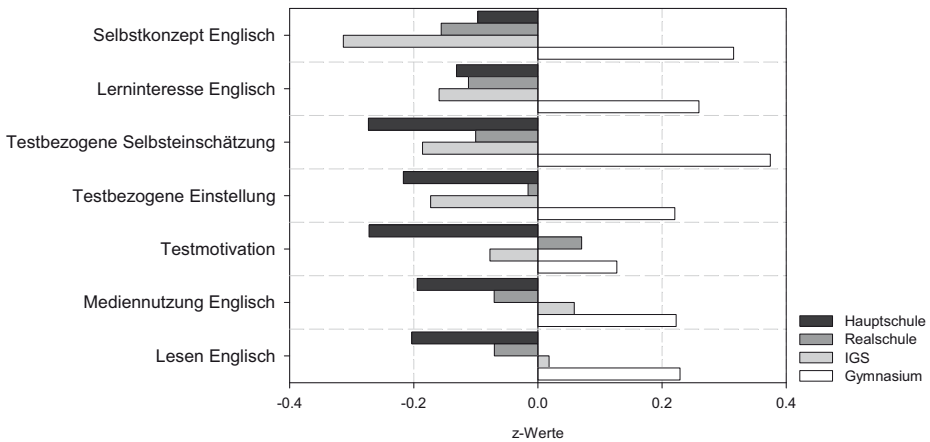


Abbildung 22.1: Bildungsgangsspezifische Motivationsprofile in Bezug auf Englisch (HS = Hauptschule, RS = Realschule, IGS = Integrierte Gesamtschule, GY = Gymnasium).

Die Ergebnisse widersprechen gelegentlich zu hörenden Einschätzungen in Bezug auf die angeblich deprimierende subjektive Lage der Hauptschülerinnen und Hauptschüler. Trotz der in den Tests im Fach Englisch festgestellten Leistungsschwächen ist die subjektive Befindlichkeit – sowohl bezogen auf das Lerninteresse als auch auf das akademische Selbstkonzept Englisch – nicht schlechter (beim Selbstkonzept sogar tendenziell eher besser) als die der Schülerinnen und Schüler in Realschule und Integrierter Gesamtschule. Aus dem Rahmen fällt lediglich das deutlich günstigere Motivationsprofil der Gymnasiasten. Dagegen zeigt sich bei den auf Englisch bezogenen Lernaktivitäten der Mediennutzung und des Englischlesens das erwartete Gefälle zwischen Hauptschule auf der einen und Gymnasium auf der anderen Seite.

Es finden sich (ohne Abbildung) nur geringe *Geschlechtsunterschiede*: Zwar weisen die Mädchen (vgl. Kap. 19) durchweg bessere Englischleistungen auf und mögen dieses Fach auch lieber als die Jungen². Im akademischen Selbstkonzept schlägt sich die Leistungsüberlegenheit der Mädchen jedoch nicht nieder (vgl. hierzu auch Abbildung 22.3). In Bezug auf die Tests schätzen die Jungen ihre Testleistungen sogar besser ein, als sie tatsächlich sind³.

Was die Rolle der *Erstsprache* anbelangt, so weisen die Schülerinnen und Schüler, die angeben, in ihrer Familie sowohl Deutsch als auch eine weitere Sprache erlernt zu haben, durchweg das günstigste Motivationsprofil auf. Es handelt sich dabei überwiegend (32%) um Kinder, bei denen ein Elternteil im Ausland, der andere in Deutschland geboren ist. Dieses Ergebnismuster ergänzt die in Kapitel 20 berich-

2 Bezogen auf das Fachinteresse liegen Mädchen zu Beginn bzw. am Ende der neunten Jahrgangsstufe jeweils etwa eine viertel Standardabweichung über den Jungen (jeweils $p < .001$).

3 Bezogen auf die subjektive Kompetenz liegen Jungen zu Beginn bzw. am Ende der neunten Jahrgangsstufe jeweils etwa eine viertel Standardabweichung über den Mädchen (jeweils $p < .001$).

teten Befunde zu den Sprachkompetenzen von Schülerinnen und Schüler mit nicht-deutscher Erstsprache.

Außerschulische Beschäftigung mit der englischen Sprache unter der Lupe

In welcher Weise beschäftigen sich Schülerinnen und Schüler im Alltag – also über die Schule und Hausaufgaben hinaus – mit der englischen Sprache? Dies zu erfahren ist deshalb wichtig, weil sich daraus möglicherweise Ansatzpunkte für eine stärker an der Lebenswelt der Schülerinnen und Schüler orientierte Motivierung im Englischunterricht ergeben könnten. In Tabelle 22.4 sind die von den Schülerinnen und Schülern angegebenen Nutzungshäufigkeiten englischsprachiger Medien aufgelistet.

Tabelle 22.4: Prozentsatz der Schülerinnen und Schüler, die angeben, mindestens ein paar Mal pro Monat englischsprachige Texte zu lesen, englischsprachige Medien zu nutzen und Mails auf Englisch zu schreiben.

	Bildungsgang			
	HS	RS	IGS	GY
Lesen von Texten				
Tageszeitungen	9	9	11	10
Comics, Bildgeschichten	17	17	20	20
Romane, Erzählungen	13	13	16	21
Sachbücher	16	13	21	10
Emails	27	33	31	40
Web-Seiten	36	49	45	61
Song-Texte	68	75	70	80
Manuale	28	29	31	34
Zeitschriften	15	14	18	19
Nutzung von Medien				
TV-Nachrichten	20	24	26	29
Dokumentarsendungen	18	19	22	20
Filme	21	19	21	27
Schreiben von E-Mails	20	23	26	33

Die Ergebnisse könnten insbesondere für die Hauptschule Hinweise geben, auf welche Weise Schülerinnen und Schüler im Englischunterricht vermehrt zum Hören und Lesen sowie zum selbstständigen Sprechen und Schreiben in englischer Sprache veranlasst werden könnten (etwa Songtexte, Webseiten, Emails und Manuale).

22.2 Motivation und Englischleistungen

Korrelative Ergebnisse

Wir betrachten zwei unterschiedliche Indikatoren der Englischkompetenz: die Testleistung und die Zeugnisnote, um so eventuellen Unterschieden im Muster der mo-

tivationalen Bedingungsfaktoren beider Leistungsindikatoren auf die Spur zu kommen; beide hängen nur mittelhoch zusammen ($r = .45$, siehe Tabelle 22.5). Hierzu werden die Partialkorrelationen zwischen motivationalen Bedingungsfaktoren beider Leistungsindikatoren berechnet, indem Mehrebenenanalysen unter Kontrolle des Sozialstatus, der kognitiven Grundfähigkeit, des Geschlechts, der Erstsprache (Deutsch vs. andere Sprache) und des Bildungsgangs (Realschule vs. andere, Gymnasium vs. andere) durchgeführt werden.

Tabelle 22.5: Partialkorrelationen zwischen motivationalen Schülermerkmalen und der globalen Englischtestleistung bzw. der (erwarteten) Englisch-Zeugnisnote am Ende des neunten Jahrgangs (Mehrebenenanalyse).

	9. Jgst.	Englisch-Testleistung			Englisch-Zeugnisnote		
		Indivi- dual- ebene	innerhalb von Klassen	zwischen Klassen	Indivi- dual- ebene	innerhalb von Klassen	zwischen Klassen
Akademisches Selbstkonzept	Beginn	.25	.49		.45	.48	.22
	Ende	.28	.54		.54	.58	.24
Lerninteresse	Beginn	.19	.35		.32	.34	.15
	Ende	.20	.37		.36	.38	.23
Subjektive Testkompetenz	Ende	.25	.38	.10	.27	.33	
Testvalenz	Ende	.17	.25	.09	.19	.21	
Englischsprachige Medien	Ende	.11	.16		.11	.14	
Englische Texte	Ende	.10	.16		.12	.15	
Testmotivation	Beginn	.07	.08				
	Ende	.07	.07				

Anmerkung: Alle aufgeführten Korrelationen sind mit mindestens $p < .01$ signifikant, nicht signifikante Korrelationen sind nicht dargestellt.

Die Höhe der Korrelationen zwischen den überdauernden Schülervariablen „Lerninteresse“ und „Akademisches Selbstkonzept“ einerseits („traits“) und der Englischtestleistung auf der Individualebene andererseits (die Spalten „Individualebene“) entspricht in etwa den Befunden, wie sie in der Forschung zum Zusammenhang zwischen Motivation und Leistung berichtet werden (vgl. Helmke/Schrader 2006). Dekomponiert man mit Hilfe mehrebenenanalytischer Methoden die Zusammenhänge auf Individualebene in Zusammenhänge innerhalb von Klassen und zwischen Klassen (die Spalten „innerhalb von Klassen“ und „zwischen Klassen“), dann zeigen sich große Unterschiede je nach Analyseebene: Der Einfluss des akademischen Selbstkonzeptes, bei dessen Genese soziale Vergleichsprozesse (mit beschränkter Reichweite, d.h. auf die eigene Klasse und nicht auf den Altersjahrgang bezogen) eine überragende Rolle spielen (Helmke 1992), wird dann sehr groß. Aber auch die situationsgebundenen Selbsteinschätzungen (subjektive Testkompetenz, Testvalenz, Testmotivation, Beschäftigung mit englischsprachigen Medien und englischen Texten) korrelieren signifikant mit der Englischtestleistung, wenngleich schwächer als die stabilen Personenmerkmale. Schulklassenunterschiede in Bezug auf die

Englisch Gesamtttestleistung zeigen sich nur bei den beiden Merkmalen, die direkt auf die durchgeführten Tests bezogen sind („states“) und vermutlich keine sozialen Vergleichsprozesse angeregt haben.

Die Zusammenhänge zwischen Selbstkonzept und Lerninteresse mit der Englischnote sind stärker als die mit der Englischtestleistung. Dies entspricht einem immer wieder gefundenen Befundmuster und erklärt sich daher, dass die Note – im Gegensatz zum Abschneiden beim Englisch-Kompetenztest – in der Klassenöffentlichkeit bekannt und von großer Bedeutung insbesondere für die Selbsteinschätzung der eigenen Sprachkompetenz ist. Bemerkenswert und erklärungsbedürftig ist, dass auch nach Ausparialisierung der wichtigsten Hintergrundvariablen die durchschnittliche Englischnote einer Klasse – nicht dagegen die durchschnittliche Englischtestleistung – noch immer signifikant mit dem durchschnittlichen Lerninteresse der Klasse zusammenhängt (letzte Spalte).

Auslandsaufenthalte und englischsprachige Kontakte

Um einen Überblick über das Vorkommen von Aufenthalten im englischsprachigen Ausland sowie über englischsprachige Kontakte zu erhalten, wurden die Schülerinnen und Schüler u.a. gefragt, ob sie a) in den vergangenen beiden Jahren im englischsprachigen Ausland *Urlaub* oder eine *Sprachreise* gemacht oder ob sie an einem entsprechenden *Austauschprogramm* teilgenommen hätten und b) wenn ja, ob sich ihr Englisch dadurch verbessert habe. Die Ergebnisse belegen, dass die verschiedenen Modi des Aufenthaltes im englischsprachigen Ausland nicht signifikant mit höheren Zuwächsen bei der Textrekonstruktion oder dem Hörverstehen einhergehen.

Darüber hinaus wurde nach aktuellen Kontakten mit Personen im englischsprachigen Ausland gefragt (per Mail, Brief etc.). Hier zeigen sich (neben dem erwartungsgemäßen Haupteffekt des Bildungsganges) statistisch signifikante Unterschiede, wie aus Tabelle 22.6 zu entnehmen ist. Insbesondere in Realschule und Gymnasium weisen Schülerinnen und Schüler mit intensiveren Sprachkontakten bessere Werte beim Hörverstehenstest auf (in der Tabelle: die Differenz zwischen „Kontakt: ja“ und „Kontakt: nein“). Dagegen sind keine Zusammenhänge mit dem Lernzuwachs nachweisbar.

Tabelle 22.6: Hörverstehen im Englischen bei Schülerinnen und Schülern in Abhängigkeit von Bildungsgang, Aufenthalt im englischsprachigen Ausland und englischsprachigen Kontakten zu Beginn und am Ende des neunten Jahrgangs⁴.

	Hauptschule		Realschule		Integrierte Gesamtschule		Gymnasium	
	Beginn	Ende	Beginn	Ende	Beginn	Ende	Beginn	Ende
Kontakt								
nein (47%)	397	413	469	495	412	448	537	573
ja (53%)	409	431	503	530	428	451	563	598
Differenz	12	18	34	35	16	3	26	25

Pfadmodelle

Alle im Folgenden berichteten Zusammenhänge sind – sofern nicht ausdrücklich etwas anderes mitgeteilt wird – um zentrale Hintergrundmerkmale bereinigte Beziehungen⁵. Berücksichtigt wurden die folgenden Hintergrundmerkmale: Bildungsgang Realschule und Gymnasium, Sozialstatus, kognitive Grundfähigkeit, Geschlecht und Erstsprache. Die bildungsgangspezifischen Pfade werden aus Gründen der Übersichtlichkeit nicht eingezeichnet.

Wechselseitige Beeinflussung von Lerninteresse und Englischleistung

Eine im schulischen Kontext, überwiegend im Fach Mathematik und für das akademische Selbstkonzept häufig untersuchte Frage (Helmke 1992; Helmke/Van Aken 1995; Marsh u.a. 2005), ist die nach der kausalen Prädominanz von Motivation und Schulleistung: Beeinflusst die Motivation (oder das Selbstkonzept) die nachfolgende Schulleistung, ist es umgekehrt, oder zeigen sich beide Wirkungsrichtungen (bidirektionale Wirkungen)?

Für das Lerninteresse zeigen die mehrebenenanalytischen Ergebnisse in Abbildung 22.2 Folgendes: Auf Individualebene gibt es neben den sehr hohen Stabilitäten über die Zeit (.60 und .81) zwar statistisch signifikante sog. *cross-lagged*-Effekte (die diagonalen Wirkungspfade) sowohl vom Lerninteresse zu Beginn der neunten Jahrgangsstufe auf das Hörverstehen am Ende als auch vom Hörverstehen am Anfang auf das Lerninteresse am Schluss; sie sind jedoch sehr gering. Anders auf *Klassenebene*: Hier findet sich ein substanzieller *cross-lagged* Pfad: Schulklassen mit höherem Lerninteresse schneiden am Ende der neunten Jahrgangsstufe beim

4 Die Unterschiede wurden mittels zweifaktorieller ANOVAs überprüft (erster Faktor jeweils ‚Bildungsgang‘, zweiter Faktor ‚Aufenthalte im englischsprachigen Ausland‘ bzw. ‚Kontakte‘; Basis: WesVar, Fay-Replicates). Signifikante Unterschiede (für die Tests zu beiden Messzeitpunkten in der neunten Jahrgangsstufe) zeigen sich – neben dem jeweils hochsignifikanten Haupteffekt für den Faktor Bildungsgang ($p < .001$) – für den Haupteffekt ‚Kontakte‘ ($p < .001$). Hier zeigen sich auch marginal signifikante Interaktionseffekte ($p < .10$).

5 Complex- und Twolevel-Modelle bei Mplus

Hörverstehenstest signifikant besser ab. Ein das Lerninteresse fördernder Unterricht schlägt sich also letztendlich auch in der Leistungsbilanz positiv nieder.

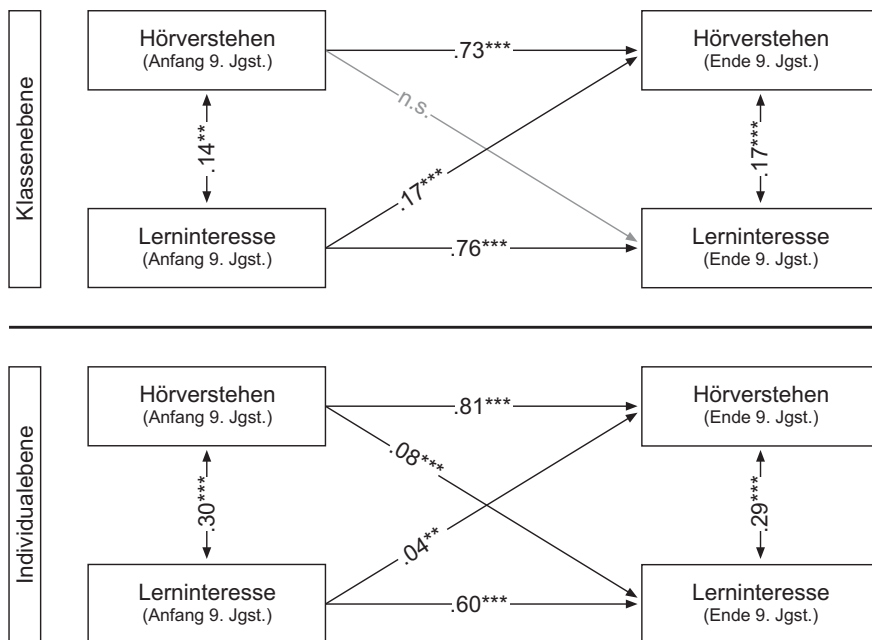


Abbildung 22.2: Dynamik der wechselseitigen Beeinflussung von Lerninteresse und Englischleistung („cross-lagged“-Modell).

Selbstkonzept und Lerninteresse als Bedingungsfaktoren der Englischleistung

In das Pfadmodell in Abbildung 22.3 gehen die beiden zentralen motivationalen Parameter, die stellvertretend für die Erwartungs- und die Wertkomponente der Motivation stehen, simultan ein. Interindividuelle Unterschiede (Geschlecht, kognitive Grundfähigkeit, Erstsprache und soziale Herkunft) stellen den Hintergrund für die Analyse der Zusammenhänge zwischen den zwei Motivationen, fachbezogenes Selbstkonzept und Lerninteresse, mit der Gesamtestleistung im Englischen am Ende der neunten Jahrgangsstufe dar. Ebenso wurde der Bildungsgang kontrolliert. Da die auf die Bildungsgänge zurückzuführenden Unterschiede bereits anderweitig dargestellt wurden, sind sie in der Abbildung 22.3 nicht aufgeführt. Es zeigt sich, dass beide motivationale Faktoren für die Englischtestleistung von Belang sind, wobei dem fachbezogenen Selbstkonzept ein stärkeres Gewicht zukommt. Damit wird ein in zahlreichen internationalen Untersuchungen zur Rolle von „academic interest“ und „academic self-concept“ gefundenes Ergebnismuster (z.B. Corbière u.a. 2006) für den Englischunterricht in Deutschland bestätigt.

Neben diesen beiden motivationalen Faktoren spielen die Hintergrundvariablen kognitive Grundfähigkeit, Geschlecht und Erstsprache eine Rolle: Schülerinnen und Schüler mit stärker ausgeprägten kognitiven Grundfähigkeiten sind erfolgreicher,

Mädchen zeigen bessere Englischleistungen als Jungen und haben ein etwas höheres Lerninteresse. Die Erstsprache hängt nur geringfügig mit dem Lerninteresse zusammen, ebenso spielt der Sozialstatus nur eine marginale Rolle.

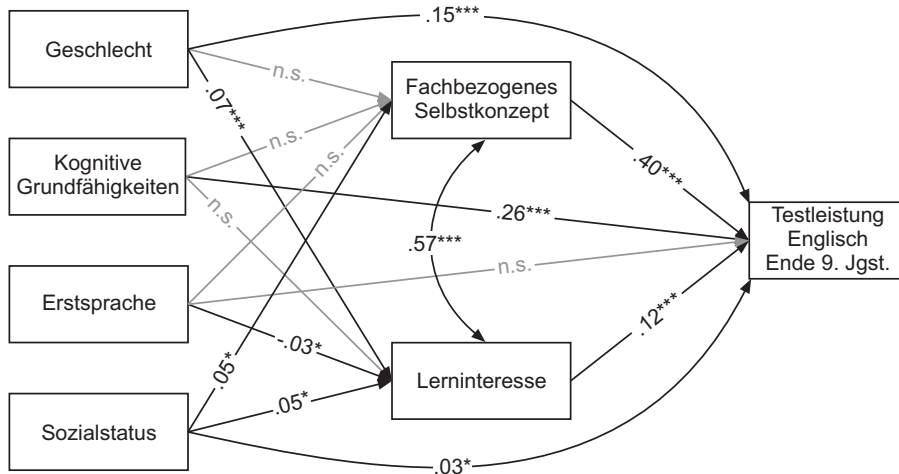


Abbildung 22.3: Pfadmodell: Motivationale Bedingungen der Englisch-Testleistung und Rolle des Hintergrundes. Bildungsgänge und die mit ihnen verknüpften Pfade sind nicht dargestellt.

22.3 Ausblick

In diesem Kapitel konnte nur ein kleiner Ausschnitt dessen berichtet werden, was im Bereich von Lernmotivation, Selbstkonzepten und Lernaktivitäten im DESI-Projekt erhoben worden ist. In weiteren und vertiefenden Analysen wird es beispielsweise um die folgenden, in diesem Kapitel nicht thematisierten Fragestellungen gehen:

- Wie *realitätsangemessen* sind Selbsteinschätzungen der eigenen Leistungen, und welche Rolle spielt das Merkmal „Realitätsangemessenheit“ im nomologischen Netzwerk von Selbstkonzept und Lerninteresse?
- Lassen sich über die hier zugrunde gelegte globale Englischleistung hinaus je nach Teilkompetenz auch differenzielle, also *kriteriumsspezifische Effekte* finden (z.B. beim Vergleich von Tests mit unterschiedlichem Anreiz und Aufforderungswert, z.B. Hörverstehen vs. Textrekonstruktion vs. Set10-Test)?
- Welche Rolle spielen die Noten in den Hauptfächern für Selbstkonzept und Motivation?
- Wie schlägt die *Sprachpraxis im Elternhaus* auf motivationale Orientierungen durch?
- Wie homogen oder heterogen sind die Motivationsprofile in den Fächern Deutsch und Englisch? Da Schüler zu beiden Bereichen Angaben gemacht haben, lassen sich intraindividuelle Profile berechnen.
- Lassen sich über das hier berichtete deskriptive Bildungsgangprofil hinaus *bildungsgangspezifische Zusammenhangsmuster* identifizieren, spielen also z.B.

Selbstkonzeptunterschiede in der Hauptschule eine ebenso große Rolle für die Leistung wie Selbstkonzeptunterschiede im Gymnasium?

- Vermittels welcher *Mechanismen* wirkt sich das akademische Selbstkonzept auf die Testleistung aus? Hier liefert die motivationspsychologische Forschungsliteratur eine Reihe ergiebiger Hinweise auf Mediationsprozesse.
- Und schließlich: Welche Rolle spielen bildungs- und schulferne *Freizeitinteressen und -aktivitäten* für die schulische Motivation und das schulbezogene Lern- und Hausaufgabenverhalten? Hier könnte die Theorie der motivationalen Konflikte von Hofer (Hofer/Saß 2006; Henn-Memmesheimer/Hofer, in Druck) den Ausgangspunkt für die Analysen darstellen.

Literatur

- Bremerich-Vos, A./Schnaitmann, G. W. (1995): Grammatik im Deutschunterricht: Leistungen, Interesse und Selbstkonzept. In: Empirische Pädagogik 9, H. 2, S. 173-191.
- Butzkamm, W. (2004): Lust zum Lehren, Lust zum Lernen. Eine neue Methodik für den Fremdsprachenunterricht. Tübingen: Francke Verlag.
- Corbière, M./Fraccaroli, F./Mbekou, V./Perron, J. (2006): Academic self-concept and academic interest measurement: A multi-sample European study. In: European Journal of Psychology of Education 21, H. 1, S. 3-15.
- Dörnyei, Z. (2000): Motivation. In: Byram, M. (Hrsg.): Routledge Encyclopedia of Language Teaching and Learning. London: Routledge, S. 425-432.
- Dörnyei, Z./Schmidt, R. (Hrsg.) (2001): *Motivation and second language acquisition*. Honolulu, HI: University of Hawaii Press.
- Fend, H. (1982): Gesamtschule im Vergleich. Weinheim: Beltz.
- Finkbeiner, C. (2003): Fremdsprachenunterricht im Sekundarbereich I. In: Bausch K.-R./Christ H./Krumm H.-J. (Hrsg.): Handbuch Fremdsprachenunterricht. Tübingen und Basel: A. Francke Verlag, S. 203-208.
- Grotjahn, R. (2003): Konzepte für die Erforschung des Lehrens und Lernens fremder Sprachen: Forschungsmethodologischer Überblick. In: Bausch K.-R./Christ H./Krumm H.-J. (Hrsg.): Handbuch Fremdsprachenunterricht. Tübingen und Basel: A. Francke Verlag, S. 493-499.
- Gruehn, S. (2000): Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Münster: Waxmann.
- Helmke, A. (1992): Selbstvertrauen und schulische Leistungen. Göttingen: Hogrefe.
- Helmke, A./Schrader, F.-W. (2001): Jenseits von TIMSS: Messung sprachlicher Leistungen, komplexe Längsschnittstudien und kulturvergleichende Analysen. Ergebnisse und Perspektiven ausgewählter Leistungsstudien. In: Weinert, F. E. (Hrsg.): Leistungsmessungen in Schulen Weinheim: Beltz, S. 237-250.
- Helmke, A./Schrader, F.-W. (2006): Determinanten der Schulleistung. In: Rost, D. H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: Psychologie Verlags Union, S. 83-94.
- Helmke, A./Van Aken, M. (1995): The causal ordering of academic achievement and self concept of ability during elementary school: A longitudinal study. In: Journal of Educational Psychology 87, H. 4, S. 624-637.
- Henn-Memmesheimer, B./Hofer, M. (in Druck): Variantenwahl, Wertorientierungen und Lernmotivation. In: Neuland, E. (Hrsg.): Variation im heutigen Deutsch: Perspektiven für den Sprachunterricht. Frankfurt/Main: Lang.
- Hofer, M./Saß, C. (2006): „Also man würde lieber rausgehen, wenn viele Hausaufgaben zu machen sind.“ In: Psychologie in Erziehung und Unterricht, H. 2, S. 122-133.

- Köller, O./Trautwein, U. (2004): Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In: Köller O./Watermann, R./Trautwein U./Lüdtke, O. (Hrsg.): Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. Opladen: Leske + Budrich, S. 285-326.
- Marsh, H. W./Trautwein, U./Lüdtke, O./Köller, O./Baumert, J. (2005): Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. In: Child Development 76, H.2, S. 397-416.
- Nold, G. (2003): The impact of the learning culture on achievement in the English as a foreign language classroom - a view from Germany. In: Rymarczyk, J./Haudeck, H. (Hrsg.): In Search of The Active Learner - Untersuchungen zu Fremdsprachenunterricht, bilingualen und interdisziplinären Kontexten. Frankfurt a. M.: Verlag Peter Lang, S. 163-184.
- Nold, G./Haudeck, H./Schnaitmann, G. W. (1997): Die Rolle von Lernstrategien im Fremdsprachenunterricht. In: Zeitschrift für Fremdsprachenforschung 8, H. 1, S. 27-50.
- Nold, G./Schnaitmann, G. W. (1995): Lernbedingungen und Lernstrategien in verschiedenen Tätigkeitsbereichen des Fremdsprachenunterrichts. In: Empirische Pädagogik 9, H. 2, S. 239-261.
- Riemer, C./Schlak, T. (2004): Der Faktor Motivation in der Fremdsprachenforschung. Einleitung in das Themenheft. In: Zeitschrift für Interkulturellen Fremdsprachenunterricht 9, H. 2, S. 3 ff. Erhältlich unter http://www.ualberta.ca/~german/ejournal/Riemer_Schlak2.htm.
- Solmecke, G. (Hrsg.) (1983): Motivation und Motivieren im Fremdsprachenunterricht. Paderborn: Schöningh.
- Solmecke, G./Boosch, A. (1981): Affektive Komponenten der Lernerpersönlichkeit und Fremdsprachenerwerb. Die Ergebnisse eines Forschungsprojekts. Tübingen: Narr.
- Weinert, F. E. (1980): Lernmotivation - Psychologische Forschung und pädagogische Aufgabe. Unterrichtswissenschaft 8, H. 3, S. 197-205.
- Wienold, G./et al. (1985): Lehrerverhalten und Lernmaterial in institutionalisierten Lehr-Lern-Prozessen. Am Beispiel des Englischanfangsunterrichtes (Band 9 - 1). Göttingen: Seminar für Wirtschaftspädagogik der Georg-August-Universität.

*Friedrich-Wilhelm Schrader / Andreas Helmke / Wolfgang Wagner /
Wolfgang Eichler / Günther Thomé / Heiner Willenberg*

23 Lernstrategien im Fach Deutsch

Lernaktivitäten umfassen ein breites Spektrum von kognitiven, metakognitiven und das Lernen unterstützenden Prozessen. Kognitive Aktivitäten betreffen die für den eigentlichen Lernvorgang maßgeblichen Prozesse (z.B. das Wiederholen, Strukturieren und Verknüpfen von Informationen), metakognitive das Planen, Überwachen und Regulieren dieser Prozesse, und ressourcenbezogene die Unterstützung des Lernprozesses durch die Beeinflussung äußerer und innerer Bedingungen (z.B. durch die Gestaltung des Arbeitsplatzes, die Steuerung der Aufmerksamkeit oder das Herstellen einer lernförderlichen Stimmung) (Friedrich/Mandl 1992; Baumert 1993; Wild 2000, 2006; Schrader/Helmke 2006). Lernstrategien sind geistige Aktivitäten, die mit dem Ziel, das Lernen und die Leistung zu verbessern, eingesetzt werden. Meistens geht man davon aus, dass es sich um bewusst kontrollierte Aktivitäten handelt. Allgemeinen Lernstrategien lassen sich bereichsspezifische Strategien gegenüberstellen, z.B. Lesestrategien zur Bewältigung von Texten. Solche Strategien sind seit langem Gegenstand fachdidaktischer Forschung (Finkbeiner 1995; Nold/Schnaitmann 1995; Willenberg 1995).

Lernstrategien gehören zu den zentralen individuellen Bedingungsfaktoren der Lernleistung (Helmke/Schrader 2001). Wegen ihrer Bedeutung für den Lernerfolg werden sie aber auch als eigenständige Zielkriterien für den Unterricht angesehen. Als solche haben sie Niederschlag gefunden in der Formulierung von Bildungsstandards für das Fach Deutsch.¹ Im Folgenden geht es aber ausschließlich um die Rolle der Lernstrategien als Determinante des Lernerfolgs und der Leistung. In der Forschungsliteratur werden nur relativ schwache – für die bereichsspezifischen Strategien meist etwas engere – Zusammenhänge zwischen Lernstrategien und der Leistung berichtet (vgl. Artelt 2000; Artelt/Demmrigh/Baumert 2001). Die Gründe dafür können theoretischer und methodischer Natur sein. Aus theoretischer Sicht spielen Lernstrategien vor allem für das selbstgesteuerte Lernen eine Rolle. Welche Bedeutung sie für das Lernen im Unterricht haben, hängt dann davon ab, inwieweit die Unterrichtsgestaltung Formen des selbstgesteuerten Lernens ermöglicht oder anregt. Aus methodischer Sicht spielt eine Rolle, dass Lernstrategien in der Regel nicht mit Kompetenzmaßen (und meistens auch nicht in der Lernsituation selbst), sondern mit Hilfe von Fragebögen erfasst werden, die auf Selbsteinschätzungen des typischen Lernverhaltens beruhen (Artelt 2000). Schneider/Schlagmüller (2002; vgl. Schlagmüller/Schneider 2007) haben stattdessen einen *Metakognitions*-Fragebogen entwickelt, der dem Kompetenzaspekt stärker Rechnung trägt.

1 <http://www.kmk.org/schul/Bildungsstandards/bildungsstandards-neu.htm>

Der Einsatz von Lernstrategien als bewusst geplante und kontrollierte Aktivitäten ist mit mentaler Anstrengung verbunden. Deshalb sind Zusammenhänge mit motivationalen Faktoren zu erwarten, wobei nicht nur das Ausmaß der Motivation, sondern auch deren Qualität von Bedeutung ist. So ist bekannt, dass intrinsische Motivation anspruchsvolle, auf eine tiefe Verarbeitung des Lernstoffs abzielende Lernaktivitäten begünstigt (Schiefele u.a. 2004; Wild 2000).

Von Lern- und metakognitiven Strategien sind mehr oder weniger gewohnheitsmäßig eingesetzte Lese-, Schreib- oder unterrichtsbezogene Aktivitäten abzugrenzen. Während Lernstrategien die Qualität des Lernverhaltens betreffen, beschreiben *Lernzeitmaße* das Lernen aus quantitativer Sicht. Gegenstand empirischer Untersuchungen sind insbesondere Hausaufgabenzeiten (Trautwein u.a. 2002).

Im vorliegenden Beitrag soll untersucht werden, welche Bedeutung verschiedene Arten von Lernstrategien und andere Aspekte des Lernverhaltens für den Lernerfolg und die Leistung im Fach Deutsch haben. Um das Bedingungsgefüge genauer aufzuhellen, werden neben der kognitiven Fähigkeit auch motivationale Merkmale der Lernenden einbezogen.

23.1 Deskriptive Ergebnisse

Angaben über deskriptive Statistiken der für dieses Kapitel relevanten Skalen und Items sind Tabelle 23.1 zu entnehmen. Nicht aufgeführt sind allgemeine Hintergrundmerkmale (Sozialstatus, Geschlecht, Erstsprache, kognitive Grundfähigkeit).

Allgemeine und bereichsspezifische Lernstrategien wurden mit aus PISA 2000 übernommenen Skalen erfasst (Antwortmodus: „nie oder fast nie, manchmal, oft, fast immer“). *Elaboration* bezeichnet das Herstellen von Verknüpfungen zwischen den zu lernenden Informationen (fünf Items; Beispiel: „Wenn ich etwas für das Fach Deutsch lerne, versuche ich, neuen Stoff mit Dingen zu verbinden, die ich bereits von anderen Fächern her kenne“). *Kontrolle* bezieht sich auf metakognitive Vorgänge des Planens, Überwachens und Regulierens von Lernvorgängen (fünf Items; Beispiel: „Wenn ich etwas für das Fach Deutsch lerne, zwingen mich zu prüfen, ob ich das Gelernte auch richtig behalten habe“). Im Unterschied zu diesen beiden bereichsübergreifenden Strategien handelt es sich bei *Lesestrategien* um bereichsspezifische Lernstrategien (sechs Items, Beispiel: „Ich streiche Textstellen an, z.B. mit einem Textmarker“). Zur Erfassung *metakognitiver Strategien* wurde der von Schneider/Schlagmüller (2002) entwickelte Metakognitions-Fragebogen eingesetzt. Die Befragten müssen für sechs Lernaufgaben jeweils fünf bis sieben Vorgehensweisen danach beurteilen, wie sehr sie das Textverständnis fördern (Beispiel: „Im Deutschunterricht fragt die Lehrkraft die Schülerinnen und Schüler, was sie tun, um einen Text möglichst gut zu verstehen. Hier sind einige Antworten, die Sie bewerten sollen: (a) Ich lese zunächst den letzten Abschnitt des Texts und gehe den Text dann von vorne durch, ... (c) Ich überfliege den Text erst einmal und konzentriere mich

dann auf das, was ich persönlich wichtig finde ...“). Die Antwortmuster werden dann danach bewertet, wie gut sie mit den Einschätzungen von Experten übereinstimmen.

Tabelle 23.1: Kennwerte der im Deutschen verwendeten Skalen und Items für die untersuchte Stichprobe (ungewichtete Angaben).

Konstrukt	Messzeit- punkt 9. Jgst.	M	SD	α	IKK ^B	Range	Stabi- lität
Lernstrategien							
Elaborationsstrategien	Beginn	2.36	0.60	.74	.04	1-4	.37
	Ende	2.21	0.65	.79	.03	1-4	
Kontrollstrategien	Beginn	2.77	0.61	.75	.04	1-4	.44
	Ende	2.57	0.65	.78	.04	1-4	
Lesestrategien	Beginn	2.52	0.56	.70	.04	1-4	.44
	Ende	2.52	0.59	.70	.05	1-4	
Metakognitive Strategien ^A	Ende	52.06	9.06	.77 ^C	.25	-3.21- 84.82	--
Lernaktivitäten							
Hausaufgabenzeit (Min.)	Ende	60.08	56.21	--	.07	0-300	--
Leseaktivitäten	Ende	1.88	0.58	.76	.05	1-4	--
Motivationale Bedingungen							
Lerninteresse	Beginn	2.67	0.75	.93	.11	1-4	.82
Fähigkeitsselbstkonzept	Beginn	3.01	0.54	.89	.07	1-4	.66
Wichtigkeit der deut- schen Sprache aus Elternsicht	Ende	3.12	0.69	.84	.08	1-4	--

Anmerkungen: ^ANach dem Messmodell von Rasch skalierte Werte (WLE-Schätzer);

^BIntraklassenkorrelation; ^CReliabilitätsschätzung aus der Raschskalierung.

Zusätzlich zu den Lernstrategien wurden folgende Lernaktivitäten erfasst: Hausaufgabenzeit („Wie viel Zeit verbringst du außerhalb des regulären Unterrichts im Durchschnitt täglich mit Hausaufgaben für Deutsch? Von Montag bis Freitag - am Wochenende: Gar keine, bis zu 15 Minuten... mehr als zwei Stunden“); beide Items wurden zu einem Gesamtwert zusammengefasst. Mit einer Skala erfasst wurden Leseaktivitäten (neun Items, Beispiel: „Wie oft liest du folgende Texte zum Vergnügen? Abenteuerromane, Kriminalromane, historische Romane“).

Als motivationale Merkmale wurden *Lerninteresse* (dreizehn Items, Beispiel: „Weil mir das Lesen Spaß macht, würde ich ungern darauf verzichten“), *Fähigkeitsselbstkonzept* (zehn Items, Beispiel: „Was wir im Deutschunterricht durchnehmen, verstehe ich meistens“) und *Wichtigkeit der deutschen Sprache aus Elternsicht* (drei Items; Beispiel: „Meine Eltern meinen, dass gutes Deutsch später sehr wichtig ist“) berücksichtigt.

Tabelle 23.2: Interkorrelationen der Merkmale (oberhalb der Hauptdiagonale: innerhalb von Klassen; unterhalb der Hauptdiagonale: zwischen Klassen). Nicht signifikante Korrelationen ($p > .05$) sind nicht dargestellt.

	Elaboration	Kontrolle	Lesestrategien	Metakognitive Strategien	Hausaufgabenzeit	Leseaktivitäten	Lerninteresse	Fähigkeits-selbstkonzept	Wichtigkeit der deutschen Spr.	Deutschnote ^A	Deutschleistung	Sozialstatus	Kognitive Grundfähigkeit	Geschlecht	Erstsprache
Elaboration		.44	n.k.		.17	.22	.11	.07	.17	.08	.04			.03	
Kontrolle	.83		.42	.15	.22	.22	.21	.19	.28	.16	.13			.18	-.03
Lesestrategien	n.k.	.65		.13	.24	.25	.23	.11	.25	.11	.07			.22	-.01
Metakognitive Strategien	.64	.56	.46				.13	.13	.15	.13	.27		.16	.19	.04
Hausaufgabenzeit	.37	.54	.31			.17	.07		.12	.06	-.08		-.06	.07	-.05
Leseaktivitäten	.97	.32	.70	.43			.36	.06	.15	.05		.07		-.07	
Lerninteresse	.62	.36	.51	.76		.71		.29	.18	.22	.33	.08	.10	.29	
Fähigkeits-selbstkonzept	.50	.46		.49			.55		.15	.40	.39	.06	.05	.18	.11
Wichtigkeit der deutschen Sprache	.67	.83	.39	.54	.25		.35	.40		.11	.16	.06		.13	
Deutschnote ^A				.21	-.33	.27	.32	.46			.36	.05	.11	.21	.09
Deutschtestleistung	.59	.61	.51	.88	-.21	.50	.91	.54	.55	.26		.08	.33	.21	.24
Sozialstatus	.60	.43	.45	.77	-.29	.59	.91	.51	.50	.37	.93		.03	-.06	.14
Kognitive Grundfähigkeit	.57	.50	.46	.84	-.21	.48	.84	.45	.49	.28	.93	.87			.05
Geschlecht (Mädchen)		.29	.25	.36			.31				.34	.22	.22		
Erstsprache (Deutsch)				.40	-.66		.30		.22	.19	.50	.52	.52	.15	

Anmerkungen: Berechnungen mit Mplus (Muthén/Muthén 2003); ^ADeutschnote wurde umgepolt; n.k. Modell konvergiert nicht.

In Tabelle 23.2 sind die auf Schulklassen- und Individualebene berechneten Korrelationen zwischen diesen Merkmalen dargestellt. Dabei ist zu berücksichtigen, dass die Intraklassen-Korrelationen (IKK) bei Elaborations-, Kontroll- und Lesestrategien nur gering sind, so dass komplexe Analysen auf Klassenebene mit diesen Merkmalen später nicht möglich sind. Da Lernstrategien individuelle Bedingungsfaktoren der Leistung sind, gilt das primäre Interesse in diesem Kapitel aber ohnehin den Analysen auf Individualebene.

Die Analysen auf Individualebene zeigen die Zusammenhänge zwischen den individuellen Schülermerkmalen, Aktivitäten und Lernstrategien, wenn die mit der unterschiedlichen Klassenzugehörigkeit der einzelnen Schüler zusammenhängenden Effekte kontrolliert werden. Die Interkorrelationen zwischen den drei mit Fragebögen erfassten Lernstrategien sind mittelhoch. Sie hängen zudem deutlich mit Hausaufgabenzeiten und Leseaktivitäten zusammen. Bedeutend schwächer sind die Zusammenhänge der drei Lernstrategien mit den metakognitiven Strategien, was

mit der unterschiedlichen Erfassung dieser Variablen zu tun haben dürfte. Lern- und metakognitive Strategien sowie Lernaktivitäten hängen auch deutlich mit motivationalen Merkmalen zusammen. Mit der kognitiven Grundfähigkeit hängen dagegen nur die metakognitiven Strategien und – schwach negativ – die Hausaufgabenzeit zusammen. Die gut belegte Leistungsüberlegenheit der Mädchen in den sprachlichen Fächern (vgl. Kapitel 19) zeigt sich ansatzweise auch im Bereich des Lernens: Mädchen weisen signifikant höhere Werte (Ausnahme: Leseaktivitäten) auf. Deutsch als Erstsprache geht mit höherer Metakognition und niedrigeren Hausaufgabenzeiten einher.

Im Unterschied zur Individualebene sind einzelne Zusammenhänge auf Klassenebene extrem hoch. Die Ergebnisse spiegeln in erheblichem Maße Bildungsgangunterschiede wider.

23.2 Lernen und schulische Leistungen

In Tabelle 23.3 sind die Zusammenhänge zwischen Lernaktivitäten und motivationalen Merkmalen mit der Leistung dargestellt, wobei zentrale Hintergrundvariablen kontrolliert sind. Hier sind vor allem die innerhalb von Klassen berechneten Ergebnisse interessant.

Tabelle 23.3: Partialkorrelationen zwischen Lernstrategien, ausgewählten Bedingungsfaktoren und der Schulleistung am Ende der neunten Jahrgangsstufe. Nicht signifikante Korrelationen ($p > .05$) sind nicht dargestellt.

Konstrukt	Deutschtestleistung			Deutschnote		
	klassen- über- greifend	inner- halb von Klassen	zwi- schen Klassen	klassen- über- greifend	inner- halb von Klassen	zwi- schen Klassen
Elaborationsstrategien	.03			.06	.08	
Kontrollstrategien	.07	.09	.07	.12	.13	
Lesestrategien	.03			.06	.07	
Metakognitive Strategien	.15	.17	.06	.05	.08	
Hausaufgabenzeit (Min)	-.03	-.06		.03	.06	-.20
Leseaktivitäten	.03	.04		.06	.15	
Lerninteresse	.15	.24		.13	.34	
Fähigkeitsselbstkonzept	.19	.32		.34	.07	.30
Wichtigkeit d. deutschen Sprache aus Elternsicht	.07	.12		.07		

Anmerkungen: Berechnungen mit Mplus (Muthén/Muthén 2003) auf der Grundlage von fünf plausible values für die Nachtestleistung. Kontrolliert wurden Sozialstatus, kognitive Grundfähigkeit, Geschlecht, Erstsprache (Deutsch vs. andere Sprache), Bildungsgang (Gymnasium vs. andere, Realschule vs. andere).

Mit der Leistung hängen vor allem die motivationalen Merkmale zusammen. Von den Maßen für das Lernverhalten korrelieren die metakognitiven Strategien am stärksten mit der Testleistung, gefolgt von den Kontrollstrategien, bei der Note

sind es die Kontrollstrategien. Dass die metakognitiven Strategien stärker mit der Testleistung als mit der Note zusammenhängen, könnte damit zu erklären sein, dass im Deutschtest auch Merkmale der Sprachbewusstheit erfasst werden. Im Folgenden wird zunächst die Rolle der metakognitiven Strategien für die Deutschtestleistung mit Hilfe eines Strukturgleichungsmodells genauer analysiert². Neben kognitiver Grundfähigkeit, Erstsprache und Geschlecht wird das Lerninteresse als motivationales Merkmal einbezogen. Angesichts der hohen Stabilität der Testleistung wird in den folgenden Analysen ausschließlich der Nachtest verwendet, ohne den Vortest einzubeziehen. Die Analysen erfolgen stets simultan auf Schulklassen- und auf Individualebene³. Dargestellt werden aber lediglich die Zusammenhänge auf Individualebene. Zur Kontrolle des Hintergrunds werden Erstsprache, Geschlecht und kognitive Grundfähigkeit sowie auf Klassenebene der Bildungsgang einbezogen.

Das in Abbildung 23.1 dargestellte Modell zeigt, dass die metakognitiven Strategien von kognitiver Fähigkeit und Motivation beeinflusst werden und sich ihrerseits günstig auf die Leistung auswirken. Schülerinnen und Schüler, die sich durch eine überdurchschnittliche Ausprägung metakognitiver Strategien auszeichnen, erzielen auch bessere Deutschleistungen. Der Zusammenhang mit der Leistung fällt für die metakognitiven Strategien fast ebenso hoch aus wie für die kognitive Grundfähigkeit oder das Lerninteresse. Metakognition hängt etwa gleich hoch mit dem Lerninteresse und der kognitiven Grundfähigkeit zusammen, wobei diese Beziehungen allerdings deutlich schwächer ausfallen als der Zusammenhang mit der Leistung. Das Ergebnismuster ist darüber hinaus vereinbar mit der Annahme, dass metakognitive Strategien den Einfluss von Motivation und kognitiver Fähigkeit auf die Leistung vermitteln. Dieses Ergebnis bleibt im Wesentlichen auch erhalten, wenn man das Modell ohne die Hintergrundmerkmale berechnet. Bei den Hintergrundmerkmalen war ursprünglich auch der Sozialstatus einbezogen worden (ohne Abbildung); er hatte aber weder auf die Leistung noch auf die metakognitiven Strategien einen signifikanten Effekt, so dass das endgültige Modell ohne dieses Merkmal gerechnet wurde. Erstsprache (Deutsch) und Geschlecht (weiblich) erweisen sich dagegen sowohl für die Leistung als auch für die metakognitiven Strategien als vorhersagekräftig: Schüler mit Deutsch als Erstsprache und Mädchen weisen höhere Werte bei der Leistung und den metakognitiven Strategien auf.

2 Wenn im Folgenden von Einflüssen oder Effekten gesprochen wird, so geschieht dies aus Gründen der sprachlichen Vereinfachung. Ein wirklicher Nachweis von Kausalzusammenhängen ist aufgrund der Untersuchungsanlage nicht möglich.

3 Die Analysen zu den Strukturgleichungsmodellen erfolgten mit dem Programm Mplus (Muthén/Muthén 2003) auf der Grundlage von fünf plausible values für die Nachtestleistung.

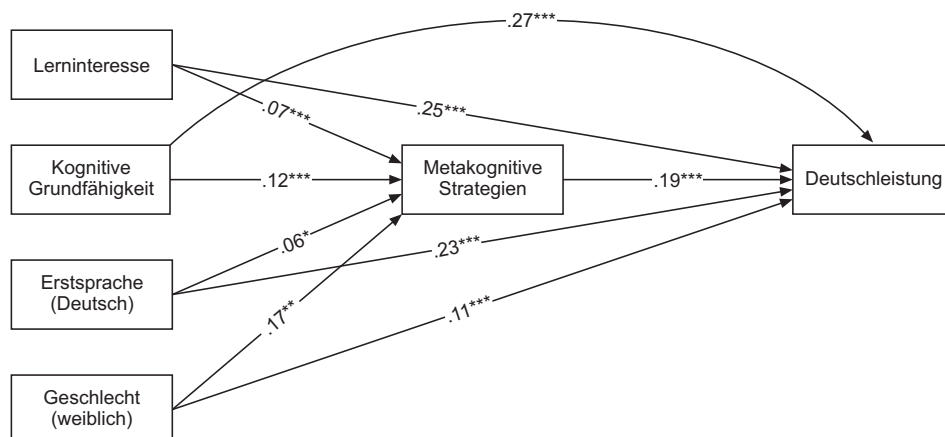


Abbildung 23.1: Pfadmodell für den Einfluss von metakognitiven Strategien auf die Deutschleistung (Individualebene) unter Berücksichtigung von Hintergrundmerkmalen.⁴

Ein ganz ähnliches Befundmuster zeigt sich auch für Kontrollstrategien (ohne Abbildung), wobei der Zusammenhang mit der Leistung für die Kontrollstrategien deutlich geringer ausfällt als für die metakognitiven Strategien. Auch lässt sich dort kein Zusammenhang zwischen kognitiver Grundfähigkeit und Lernstrategien mehr feststellen. Vergleichbare Modelle für Elaborations- und Lesestrategien (ohne Abbildung) zeigen, dass diese beiden Lernstrategien nicht mit der Leistung in Beziehung stehen, was schon aufgrund der Korrelationsergebnisse zu erwarten war.

Die Zusammenhänge mit der Hausaufgabenzeit als quantitativem Parameter des Lernverhaltens sind in Abbildung 23.2 darstellt. Der Doppelcharakter, den man diesem Merkmal auf Individualebene zuschreibt, wird auch in den Ergebnissen sichtbar: Die individuell vom Schüler aufgewendete Hausaufgabenzeit hängt positiv mit der Motivation (je höher die Motivation, umso mehr Zeit wird aufgewendet) und negativ mit der kognitiven Grundfähigkeit (je niedriger die Fähigkeit, umso mehr Zeit wird benötigt) zusammen. Dass die individuell für Hausaufgaben aufgewendete Zeit mit der Leistung negativ zusammenhängt, dürfte damit zu erklären sein, dass leistungsschwächere Schüler mehr Zeit für Hausaufgaben benötigen. Mädchen investieren mehr Zeit in Hausaufgaben als Jungen. Schüler mit nicht deutscher Sprachherkunft wenden weniger Zeit für Hausaufgaben auf als Schüler mit deutscher Sprachherkunft.

4 Da ein vollständiges Modell mit manifesten Variablen analysiert wurde, liegt ein perfekter Fit vor. Erklärte Varianz auf Individualebene: R^2 (Leistung) = .31; R^2 (Metakognitive Strategien) = .06; Klassenebene: R^2 (Leistung) = .98; R^2 (Metakognitive Strategien) = .80.

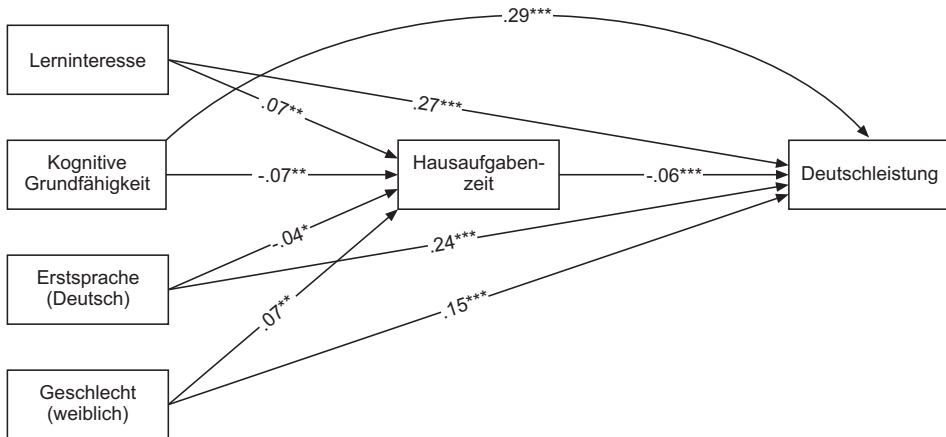


Abbildung 23.2: Pfadmodell für den Einfluss der Hausaufgabenzeit auf die Deutschleistung (Individualebene) unter Berücksichtigung von Hintergrundmerkmalen.⁵

Auf der Schulklassenebene hat die Hausaufgabenzeit eine andere Bedeutung als auf Individualebene: Sie ist hier Indikator für das Anforderungsniveau des Unterrichts (je höher die von der Klasse im Durchschnitt berichtete Hausaufgabenzeit, umso höher die gestellten Anforderungen).

Bislang war bei den Analysen der anfängliche Leistungsstand im Fach Deutsch ausgeblendet worden, der seinerseits von Merkmalen wie kognitiver Fähigkeit, Erstsprache und Geschlecht abhängen dürfte. In Abbildung 23.3 wird dargestellt, welche Rolle der Leistungsstand im Fach Deutsch zu Beginn der neunten Jahrgangsstufe spielt, wenn dieser um die genannten Einflussfaktoren bereinigt wird.

Als Maß für den Leistungsstand wird die Deutschnote zu Beginn der neunten Jahrgangsstufe verwendet, die ein grober Indikator für das Leistungsniveau in diesem Fach ist und aufgrund der nicht übermäßig hohen Korrelation mit der Nachtestleistung noch genügend Spielraum für andere Einflussfaktoren lässt. Führt man die Analyse wie bislang auch auf zwei Ebenen durch, dann bedeutet dies, dass auf Individualebene die um Klasseneffekte bereinigten Noten eingehen. Der Leistungsstand zu Beginn hängt erwartungsgemäß deutlich mit der Deutschleistung am Ende der neunten Jahrgangsstufe zusammen, aber auch mit den Kontrollstrategien. Schülerinnen und Schüler mit Deutsch als Erstsprache weisen ebenso wie Mädchen sowohl signifikant bessere Deutschnoten zu Beginn als auch bessere Deutschtestleistungen am Ende der neunten Jahrgangsstufe auf. Deutschnote und Deutschtestleistung hängen darüber hinaus signifikant positiv mit der kognitiven Grundfähigkeit zusammen. Erstsprache und kognitive Fähigkeit haben dagegen keinen direkten Effekt auf die Kontrollstrategien.

⁵ Perfekter Modell-Fit. Within-Ebene: R^2 (Leistung) = .27; R^2 (Hausaufgabenzeit) = .02; Erklärte Varianz auf Individualebene: R^2 (Leistung) = .95; R^2 (Hausaufgabenzeit) = .38

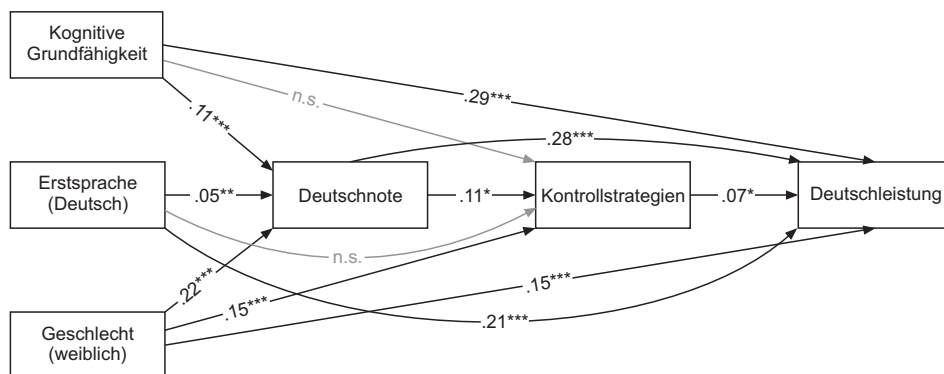


Abbildung 23.3: Pfadmodell für den Einfluss von Kontrollstrategien und Leistungsstand auf die Deutschleistung (Individualebene) unter Berücksichtigung von Hintergrundmerkmalen.⁶

23.3 Profile des Lernverhaltens

Weitere Aufschlüsse zur Bedeutung der Lernstrategien lassen sich gewinnen, wenn man von einer variablen- zu einer personenorientierten Betrachtungsweise übergeht und Personen vergleicht, die hinsichtlich aller vier Strategien entweder günstige oder ungünstige Ausprägungen aufweisen oder eine Mittelposition einnehmen.

Zu diesem Zweck wurden die Schülerinnen und Schüler, getrennt nach Bildungsgang, für jedes der vier Lernstrategiemerkmale in drei Gruppen (niedrige, mittlere, hohe Ausprägung) eingeteilt. Diejenigen Schülerinnen und Schüler, die bei allen vier Merkmalen zur oberen oder unteren Gruppe gehören, wurden jeweils zusammengefasst und der verbleibenden Restgruppe der Schüler gegenübergestellt. Diese drei Schülergruppen werden dann im Hinblick auf verschiedene lern- und leistungsrelevante Merkmale verglichen. Die für die vier Bildungsgänge resultierenden Profile sind in Abbildung 23.4 dargestellt.

⁶ Perfekter Modell-Fit. Erklärte Varianz auf Individualebene: R^2 (Leistung) = .29; R^2 (Note) = .06; R^2 (Kontrollstrategien) = .04; Klassenebene: R^2 (Leistung) = .91; R^2 (Note) = .25

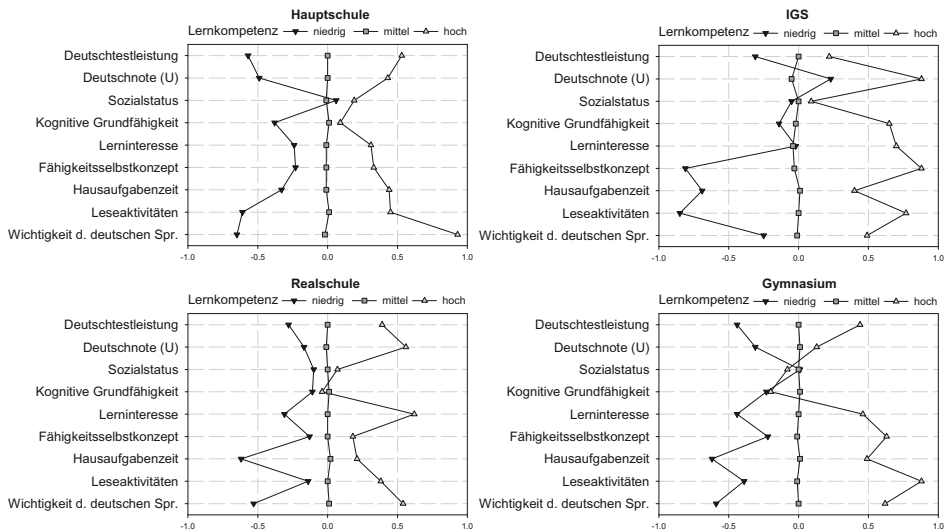


Abbildung 23.4: Merkmalsprofile für Schüler mit niedriger, mittlerer und hoher Lernkompetenz, separat nach Bildungsgang (Profilmerkmale klassenweise standardisiert).

23.4 Diskussion

Die Ergebnisse zeigen, dass Kontrollstrategien und metakognitive Strategien auch unabhängig von anderen Personenmerkmalen mit der am Ende der neunten Jahrgangsstufe gemessenen Leistung im Fach Deutsch zusammenhängen. Die Effekte sind zwar nicht groß, aber signifikant. Bemerkenswert ist dies bei den Kontrollstrategien, da diese mit einem Fragebogen erfasst wurden und damit eher Selbsteinschätzungen des habituellen Lernverhaltens als tatsächlich gemessene Kompetenzen repräsentieren. Auch die niedrigen und in den Pfadanalysen nicht mehr nachweisbaren Zusammenhänge mit der kognitiven Grundfähigkeit lassen es fraglich erscheinen, dass tatsächlich eine Kompetenz erfasst wurde. Anders bei den durch den Vergleich mit Expertenurteilen erfassten metakognitiven Strategien. Hier gibt es deutliche Zusammenhänge mit der kognitiven Grundfähigkeit, die auch in den Pfadanalysen erhalten bleiben.

Die Interpretation der Zusammenhänge erfolgte unter der Annahme, dass Merkmale des Lernverhaltens einen Einfluss auf die Leistung haben. Es muss aber betont werden, dass das Untersuchungsdesign keine definitive Absicherung von Kausaleinflüssen erlaubt. Dafür wäre ein experimentelles Vorgehen, z.B. durch Trainingsstudien, unerlässlich. Lernstrategien als bewusst und kontrolliert eingesetzte Vorgehensweisen erfordern im Unterschied zu nicht-strategischen Aktivitäten eine erhöhte mentale Anstrengung. Die deshalb zu erwartenden Zusammenhänge mit motivationalen Merkmalen lassen sich in der Tat belegen. Lernzeitmaße bilden das

Lernverhalten aus einer anderen Perspektive ab. Im vorliegenden Beitrag wurden Beziehungen zwischen der individuell aufgewendeten Hausaufgabenzeit und der Leistung analysiert. Der gefundene Zusammenhang ist negativ, was darauf hindeutet, dass schwächere Schüler längere Hausaufgabenzeiten benötigen. Das dargestellte Pfadmodell bestätigt den erwarteten Zusammenhang, wonach die in häusliches Lernen investierte Zeit positiv mit der Motivation, aber negativ mit der Fähigkeit zusammenhängt. Die Gegenüberstellung unterschiedlicher Lernergruppen, die sich aus der Kombination der einzelnen Lernstrategiemaße definieren lassen, zeigt charakteristische Unterschiede in der Leistung, der Motivation und den Lernaktivitäten, nicht aber im Hinblick auf Sozialstatus und kognitive Grundfähigkeit.

Insgesamt gesehen zeigen die Analysen, dass Merkmale des Lernverhaltens für den Lernerfolg und die Leistung im Fach Deutsch von Bedeutung sind. Dies unterstreicht, dass es sinnvoll ist, das Lernverhalten selbst als ein wichtiges Zielkriterium des schulischen Unterrichts anzusehen.

Literatur

- Artelt, C. (2000): Strategisches Lernen. Münster: Waxmann.
- Artelt, C./Demmrich, A./Baumert, J. (2001): Selbstreguliertes Lernen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich, S. 271-298.
- Baumert, J. (1993): Lernstrategien, motivationale Orientierungen und Selbstwirksamkeitsüberzeugungen im Kontext schulischen Lernens. In: Unterrichtswissenschaft 21, S. 327-354.
- Finkbeiner, C. (1995): Zur Erhebung von textverstehensrelevanten Lernstrategien und Interessen im Fremdsprachenunterricht. Entwicklung zweier Fragebögen. Empirische Pädagogik 9, S. 193-219.
- Friedrich, H.F./Mandl, H. (1992): Lern- und Denkstrategien – Ein Problemaufriß. In: Mandl, H./Friedrich, H.F. (Hrsg.): Lern- und Denkstrategien. Analyse und Intervention. Göttingen: Hogrefe, S. 3-54.
- Helmke, A./Schrader, F.-W. (2001): School achievement, cognitive and motivational determinants. In: Smelser, N.J./Baltes, P.B. (Hrsg.): International encyclopedia of the social and behavioral sciences (Bd. 20). Oxford: Pergamon, S. 13552-13556.
- Muthén, L.K./Muthén, B.O. (2004). Mplus User's Guide. Statistical Analysis With Latent Variables. Los Angeles, CA: Muthén & Muthén.
- Nold, G./Schnaitmann, G.W. (1995): Lernbedingungen und Lernstrategien in verschiedenen Tätigkeitsbereichen des Fremdsprachenunterrichts. In: Empirische Pädagogik 9, S. 239-261.
- Schiefele, U./Artelt, C./Schneider, W./Stanat, P. (Hrsg.) (2004): Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schlagmüller, M./Schneider, W. (2007): WLST 7-12. Würzburger Lesestrategie-Wissenstest für die Klassen 7-12. Göttingen: Hogrefe.
- Schneider, W./Schlagmüller, M. (2002): Metakognitives Wissen über Textverarbeitung. Würzburg: Lehrstuhl für Psychologie IV, Universität Würzburg.
- Schrader, F.-W./Helmke, A. (2006): Lernstrategien. In: Brockhaus Enzyklopädie. (Band 16). Leipzig: Brockhaus, S. 638-639.
- Trautwein, U./Köller, O./Schmitz, B./Baumert, J. (2002): Do homework assignments enhance achievement? A multilevel analysis in 7th-grade mathematics. In: Contemporary Educational Psychology 27, S. 26-50.

Wild, K.-P. (2000): Lernstrategien im Studium. Münster: Waxmann.

Wild, K.-P. (2006): Lernstrategien und Lernstile. In: Rost, D.H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: Psychologie Verlags Union, S. 427-432.

Willenberg, H. (1995): Die Strategien des Lesens und Lernens sind individuell gemischt. In: Empirische Pädagogik 9, S. 263-283.

*Friedrich-Wilhelm Schrader / Andreas Helmke / Wolfgang Wagner /
Günter Nold / Konrad Schröder*

24 Lernstrategien im Fach Englisch

Lernstrategien sind geistige Aktivitäten, die mit dem Ziel des Lernens, Verstehens und Wissenserwerbs mehr oder weniger bewusst eingesetzt werden. Neben allgemeinen Lernstrategien, die grundlegende kognitive, metakognitive und ressourcenbezogene Aktivitäten umfassen (zum allgemeinen Hintergrund vgl. Kapitel 23), spielen bereichsspezifische Lernstrategien eine wichtige Rolle. Dies gilt in besonderem Maße für das Lernen von Fremdsprachen, das in den letzten Jahren zum Gegenstand der Lernstrategie-Forschung geworden ist (Bimmel 2006; Finkbeiner 1995; Haudeck 1998; Nold 1992, 2000, 2003; Nold/Haudeck/Schnaitmann 1997; Nold/Schnaitmann 1995; Raupach 2006; Wolff 1998; Zimmermann/Plessner 1998). Für das Fremdsprachenlernen sinnvolle Strategien umfassen unter anderem gezielte Bemühungen, die Fremdsprache möglichst oft und systematisch einzusetzen, mit fehlenden Informationen (z.B. nicht bekannten Bedeutungen) umzugehen oder Lerngelegenheiten zu nutzen. Diese Strategien sind Gegenstand des Strategy Inventory for Language Learning (SILL) (Oxford/Burry-Stock 1995), einem der bekanntesten Fragebogeninstrumente in diesem Bereich.

Mit Lernstrategien verbindet man gewöhnlich die Erwartung, dass sie die Qualität des Lernens und der Leistung beeinflussen. Es gibt allerdings noch zu wenig empirisch gesichertes Wissen dazu, ob und in welchem Umfang allgemeine und spezifische Lernstrategien zur Vorhersage und Erklärung von fremdsprachlichen Leistungen beitragen. Für Deutschland liegen aus den Pilotierungsuntersuchungen zu DESI erste Ergebnisse zu den mit dem SILL erfassten Lernstrategien vor (Schrader u.a. 2003). Dabei konnten Zusammenhänge mit der Englischnote belegt werden.

Neben Lernstrategien sind auch nicht-strategische, d.h. meist nicht mit gezielter Lernabsicht eingesetzte Lernaktivitäten wie etwa das Lesen und die Mediennutzung in der Fremdsprache wichtig. Der häufige Einsatz solcher Aktivitäten sollte längerfristig zu einer Verbesserung fremdsprachlicher Leistungen führen. Diese quantitative Sicht auf das Lernverhalten kommt noch stärker zum Tragen, wenn Lernzeitmaße, also die für bestimmte Aktivitäten aufgewendete Zeit, berücksichtigt werden. Insbesondere Untersuchungen zur Hausaufgabenzeit haben hier eine eigenständige Forschungstradition begründet (vgl. etwa Trautwein u.a. 2002).

Grundsätzlich ist davon auszugehen, dass Quantität und Qualität des Lernens sowohl von motivationalen Faktoren als auch von kognitiven Fähigkeiten abhängen. Im vorliegenden Beitrag soll deshalb untersucht werden, welche Bedeutung allgemeine und sprachspezifische Lernstrategien für das Lernen und die Leistung haben und welche Bedingungen dabei eine Rolle spielen. Es soll ferner geprüft werden, welche Bedeutung motivationale Merkmale über Hintergrundmerkmale wie Bildungsgang,

kognitive Grundfähigkeit, Sozialstatus, Erstsprache und Geschlecht hinaus für das Lernen und die Leistung haben.

24.1 Deskriptive Ergebnisse

Angaben über statistische Kennwerte für die verwendeten Variablen sind Tabelle 24.1 zu entnehmen. Nicht aufgeführt sind allgemeine Hintergrundmerkmale (Sozialstatus, Geschlecht, Erstsprache, kognitive Grundfähigkeit).

Tabelle 24.1: Kennwerte der im Englischen verwendeten Skalen und Items für die untersuchte Stichprobe (ungewichtete Angaben).

Konstrukt	Messzeit- punkt Jahrgangs- stufe 9	M	SD	α	IKK ^A	Range	Stabi- lität
Lernstrategien							
Elaborationsstrategien	Beginn	2.30	0.63	.74	.02	1-4	.42
	Ende	2.25	0.63	.76	.03	1-4	
Kontrollstrategien	Beginn	2.80	0.64	.77	.05	1-4	.43
	Ende	2.61	0.64	.76	.04	1-4	
Metakognition ^B	Ende	52.04	9.45	.77 ^C	.27	-3.21- 84.82	--
Sprachnutzung	Beginn	1.94	0.70	.85	.04	1-4	.61
	Ende	1.98	0.72	.86	.05	1-4	
Umgang mit feh- lender Information	Beginn	2.24	0.59	.64	.09	1-4	.42
	Ende	2.28	0.61	.65	.10	1-4	
Aufsuchen von Lerngelegenheiten	Beginn	2.40	0.68	.76	.05	1-4	.46
	Ende	2.34	0.66	.75	.03	1-4	
Lernaktivitäten							
Hausaufgabenzeit (Min.)	Ende	58.03	53.66	--	.04	0-300	--
Leseaktivitäten	Beginn	1.95	0.57	.78	.09	1-4	.59
	Ende	1.97	0.61	.82	.06	1-4	
Mediennutzung	Beginn	1.85	0.74	.75	.05	1-4	--
Motivationale Bedingungen							
Lerninteresse	Beginn	2.58	0.65	.77	.09	1-4	--
Fähigkeitsselbst- konzept	Beginn	2.98	0.62	.92	.06	1-4	--
Wertschätzung des Fachs	Ende	3.27	0.72	.87	.14	1-4	--

Anmerkungen: ^AIntraklassenkorrelation; ^BNach dem Messmodell von Rasch skalierte Werte (WLE-Schätzer); ^CReliabilitätsschätzung aus der Raschskalierung.

Die beiden allgemeinen Lernstrategien *Elaboration* und *Kontrolle* wurden mit zwei aus PISA 2000 übernommenen Skalen erfasst, *metakognitive Strategien* mit dem von Schneider/Schlagmüller (2002) entwickelten Instrument (für Einzelheiten

vgl. Kapitel 23). Für die Messung bereichsspezifischer Lernstrategien wurden drei Skalen des SILL (Oxford/Burry-Stock 1995) in einer etwas gekürzten Fassung eingesetzt: *Sprachnutzung* (sechs Items; Beispiel: „Ich versuche, wie ein englischer Muttersprachler zu sprechen“); *Umgang mit fehlender Information* (fünf Items; Beispiel: „Ich finde die Bedeutung eines englischen Wortes dadurch, dass ich es in Teile zergliedere, die ich verstehe“); *Aufsuchen von Lerngelegenheiten* (vier Items; Beispiel: „Ich achte darauf, welche Fehler ich in Englisch mache, und nutze diese Information, um mich zu verbessern“). Antwortmöglichkeiten (für die Fragebogenskalen): *nie oder fast nie, manchmal, oft, fast immer*.

Lernaktivitäten. Zusätzlich zu den Lernstrategien wurden folgende lernbezogene Aktivitäten erfasst: *Hausaufgabenzeit* („Wie viel Zeit verbringst du außerhalb des regulären Unterrichts im Durchschnitt täglich mit Hausaufgaben für Englisch? Von Montag bis Freitag – am Wochenende: Gar keine, bis zu 15 Minuten... mehr als zwei Stunden“); beide Zeitangaben wurden zu einem Gesamtwert zusammengefasst. Mit jeweils einer Skala erfasst wurden: *Leseaktivitäten* in englischer Sprache („Wie oft liest du zu deinem Vergnügen?“; neun Items, Beispiele: „Tageszeitungen [in englischer Sprache] ... Romane, Erzählungen, Geschichten [in englischer Sprache]“), *Mediennutzung* in englischer Sprache (vier Items; Beispiel: „Wie oft siehst du englischsprachige TV-Nachrichten (z.B. CNN)?“) und *Wertschätzung des Fachs Englisch in der Familie* (drei Items; Beispiel: „Meine Eltern meinen, dass ich Englischkenntnisse später gut gebrauchen kann“).

In Tabelle 24.2 sind die auf Schulklassen- und Individualebene berechneten Interkorrelationen dieser Merkmale dargestellt. Die mit einem Fragebogen erfassten Lernstrategien hängen auf Individualebene deutlich zusammen, während die Korrelationen mit der Metakognition (möglicherweise aufgrund der andersartigen Erfassung) recht niedrig sind. Die Lernstrategien hängen mäßig bis mittelhoch mit den nicht-strategischen Lernaktivitäten (Ausnahme: Metakognition) und den motivationalen Merkmalen zusammen. Die Zusammenhänge auf Schulklassenebene fallen teilweise recht hoch aus, was auch damit zu tun haben dürfte, dass der Bildungsgang hier nicht kontrolliert wurde.

Tabelle 24.2: Interkorrelationen der verwendeten Merkmale (oberhalb der Hauptdiagonale: innerhalb von Klassen; unterhalb der Hauptdiagonale: zwischen Klassen). Nicht signifikante Korrelationen ($p > .05$) sind nicht dargestellt.

	Elaboration	Kontrolle	Sprachnutzung	Umgang mit fehlender Inf.	Lerngelegenheiten	Metakognition	Hausaufgabenzeit	Leseaktivitäten	Mediennutzung	Lerninteresse	Fähigkeits-selbstkonzept	Wertschätzung des Fachs	Englischnote ^A	Englishtestleistung	Sozialstatus	Kognitive Grundfähigkeit	Geschlecht (weiblich)	Erstsprache (Deutsch)
Elaboration		.40	.34	.38	.40	.05	.16	.16	.14	.16	.09	.12	.10	.11	.06		.04	.06
Kontrolle	.66		.27	.28	.46	.16	.19	.10	.06	.24	.17	.22	.14	.12			.17	
Sprachnutzung	.88	.36		.48	.58		.15	.40	.39	.43	.29	.12	.22	.24	.08		.04	
Umgang mit fehlender Information	.85	.80	.79		.44	.08	.09	.20	.19	.19	.14	.11	.16	.22	.05	.07	.04	
Lerngelegenheiten	.95	.99	.90	.89		.08	.20	.25	.21	.34	.21	.20	.16	.15	.05		.09	
Metakognition	.51	.53		.67	.37						.06	.17	.10	.20		.15	.19	.03
Hausaufgabenzeit		.37		.34	.40	.26		.07	.07	.08	-.06		-.04	-.06			.07	
Leseaktivitäten	.80	.41	.64	.86	.70	.70			.64	.36	.22	.09	.12	.14	.08			
Mediennutzung	.84	.46	.55	.84	.81	.56		.92		.33	.19	.06	.09	.13	.07	-.04	-.08	
Lerninteresse	.67	.34	.71	.61	.64	.55		.81	.69		.60	.17	.36	.35	.04		.07	
Fähigkeits-selbstkonzept	.61		.43	.49		.46		.58	.57	.84		.15	.51	.50	.05			
Wertschätzung des Fachs	.74	.79	.29	.75	.61	.89	.32	.65	.59	.52	.34		.11	.17	.06	.03	.16	
Englisch-note ^A	.27			.17		.20	-.27	.17		.40	.57			.46		.09	.11	
Englishtestleistung	.76	.60	.54	.87	.59	.83	.21	.82	.69	.62	.51	.83	.22		.05	.26	.14	
Sozialstatus	.75	.46	.51	.83	.43	.76		.73	.58	.56	.53	.77	.32	.92			-.06	.15
Kognitive Grundfähigkeit	.64	.47	.34	.72	.35	.87		.72	.64	.55	.44	.82	.22	.88	.86			.05
Geschlecht (weiblich)		.32		.21		.39	.24	.23			.24	.37		.31	.23	.26		
Erstsprache (Deutsch)			-.18	.18	-.26	.44	-.26					.29	.14	.38	.56	.52	.19	

Anmerkungen: Berechnungen mit Mplus (Muthén/Muthén 2004). ^AEnglischnote wurde umgepolt.

24.2 Lernen und schulische Leistungen

In Tabelle 24.3 sind die Zusammenhänge der Lernaktivitäten sowie der motivationalen Merkmale mit der Leistung dargestellt, wobei zentrale Hintergrundvariablen

kontrolliert sind. Hier sind vor allem die innerhalb der Klassen bestehenden Zusammenhänge interessant.

Tabelle 24.3: Partialkorrelationen zwischen Lernstrategien, ausgewählten Bedingungsfaktoren und der Schulleistung am Ende der neunten Jahrgangsstufe. Kontrolliert wurden Sozialstatus, kognitive Grundfähigkeit, Geschlecht, Erstsprache (Deutsch vs. andere Sprache), Bildungsgang (Gymnasium vs. andere, Realschule vs. andere) auf beiden Ebenen. Nicht signifikante Korrelationen ($p > .05$) sind nicht dargestellt.

Konstrukt	Englishtestleistung			Englischnote		
	klassen- über- greifend	inner- halb von Klassen	zwi- schen Klassen	klassen- über- greifend	inner- halb von Klassen	zwi- schen Klassen
Elaborationsstrategien	.05	.09		.08	.11	
Kontrollstrategien	.08	.10		.10	.13	
Sprachnutzung	.14	.24	.10	.19	.22	
Umgang mit fehlender Information	.14	.20		.13	.15	
Lerngelegenheiten	.10	.14		.13	.16	
Metakognition	.11	.13	.05	.05	.07	
Hausaufgabenzeit		-.06		-.06		-.24
Leseaktivitäten	.11	.16		.10	.12	
Mediennutzung	.10	.15		.08	.10	
Lerninteresse	.19	.35		.33	.35	.17
Fähigkeitsselbstkonzept	.25	.51		.49	.51	.35
Wertschätzung des Fachs	.10	.15		.06	.10	-.11

Anmerkungen: Berechnungen mit Mplus (Muthén/Muthén 2004) auf der Grundlage von fünf plausible values für die Nachtestleistung.

Im Folgenden werden über diese einfachen Zusammenhänge hinaus komplexe Beziehungen mit Hilfe von Strukturgleichungsmodellen analysiert. Angesichts der hohen Stabilität der Testleistungen wird für die Analysen ausschließlich der Posttest verwendet, ohne den Vortest einzubeziehen. Als allgemeine Bedingungsfaktoren werden jeweils kognitive Grundfähigkeit, Lerninteresse, Sozialstatus und Erstsprache berücksichtigt. Die Analysen erfolgen stets simultan auf Schulklassen- und auf Individualebene unter Kontrolle des Bildungsgangs.

Zunächst werden die Beziehungen zwischen metakognitiven Strategien und der Leistung analysiert¹. Das in Abbildung 24.1 dargestellte Modell zeigt, dass Schülerinnen und Schüler mit höherer Ausprägung metakognitiver Strategien auch bessere Leistungen im Englischtest erzielen. Stärker noch als von den metakognitiven Strategien wird die Englischleistung allerdings durch die kognitive Grundfähigkeit und das Lerninteresse vorhergesagt. Mädchen erzielen deutlich bessere Leistungen

¹ Wenn im Folgenden von Einflüssen oder Effekten gesprochen wird, so geschieht dies aus Gründen der sprachlichen Vereinfachung. Ein wirklicher Nachweis von Kausalzusammenhängen ist aufgrund der Untersuchungsanlage nicht möglich.

als Jungen. Der Sozialstatus hat ebenfalls einen Einfluss auf die Leistung, nicht dagegen die Erstsprache. Mädchen zeichnen sich durch eine höhere Ausprägung metakognitiver Strategien aus als Jungen. Darüber hinaus zeigt nur noch die kognitive Grundfähigkeit einen Effekt auf die Metakognition.

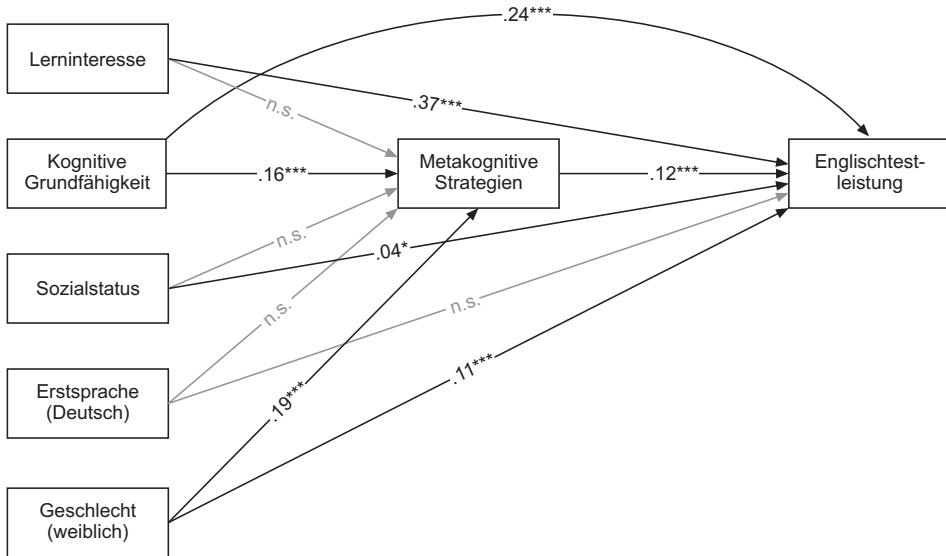


Abbildung 24.1: Pfadmodell für den Einfluss von metakognitiven Strategien auf die Englischtestleistung (Individualebene) unter Berücksichtigung von Hintergrundmerkmalen.²

Auf Klassenebene (ohne Abbildung) spielen neben dem Niveau der Metakognition und der Höhe des Sozialstatus (je höher beide sind, umso besser die Leistung) auch der Anteil der Schülerinnen und Schüler mit Deutsch als Erstsprache (je niedriger dieser Anteil, umso höher die Leistung) und der Mädchenanteil (je höher dieser ist, umso höher die Leistung) eine Rolle. Für die Vorhersage des Klassenniveaus bei den metakognitiven Strategien erweisen sich hier lediglich das Niveau der kognitiven Grundfähigkeit und der Mädchenanteil als relevant (je höher beide sind, umso günstiger für die Metakognition). Bildungsgangsunterschiede in der Englischtestleistung sind nicht mehr nachweisbar, wenn man Sozialstatus, Erstsprache, Geschlecht, kognitive Grundfähigkeit und Lerninteresse als Hintergrund- bzw. Erklärungsmerkmale verwendet.

Während metakognitive Strategien die Qualität des Lernens betreffen, charakterisiert die Hausaufgabenzeit das Lernen aus quantitativer Sicht. Wie Abbildung 24.2 zeigt, hat sie auf Individualebene einen negativen Zusammenhang mit der Englischleistung: Je größer die vom einzelnen Schüler aufgewendete Hausaufgabenzeit, umso niedriger die Leistung. Dies deutet darauf hin, dass leistungs-

2 Da ein vollständiges Modell mit manifesten Variablen analysiert wurde, liegt ein perfekter Fit vor. Erklärte Varianz auf Individualebene: $R^2 = .24$ für Leistung, $R^2 = .07$ für metakognitive Strategien; auf Klassenebene: $R^2 = .94$ für Leistung, $R^2 = .73$ für metakognitive Strategien.

schwache Schüler mehr Zeit für die Hausaufgaben benötigen. Die individuell aufgewendete Hausaufgabenzeit hängt nicht mit der kognitiven Grundfähigkeit, wohl aber positiv mit dem Lerninteresse und der Erstsprache zusammen. Lerninteresse, kognitive Grundfähigkeit und Sozialstatus haben auch hier wieder einen positiven Effekt auf die Leistung. Mädchen wenden mehr Zeit für Hausaufgaben auf und erzielen auch bessere Leistungen als Jungen.

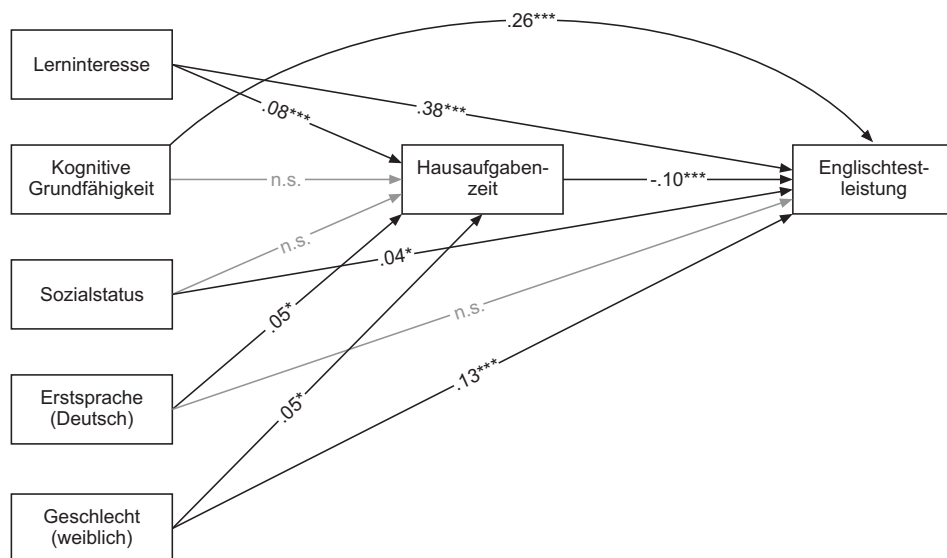


Abbildung 24.2: Pfadmodell für den Einfluss der Hausaufgabenzeit auf die Englischtestleistung (Individualebene) unter Berücksichtigung von Hintergrundmerkmalen.³

Auf Zusammenhänge auf Klassenebene soll hier nicht eingegangen werden, da diese wegen der niedrigen Intraklassenkorrelation der Hausaufgabenzeit nicht aussagekräftig sind.

Als drittes Merkmal sollen sprachspezifische Lernstrategien betrachtet werden. Wie aus Tabelle 24.3 hervorgeht, weisen innerhalb der Klassen Sprachnutzung und Umgang mit fehlender Information die höchsten Zusammenhänge mit der Testleistung auf. Da das zuletzt genannte Merkmal wegen der höheren Intraklassenkorrelation für Analysen ergiebiger erscheint, sind in Abbildung 24.3 Ergebnisse zu diesem Merkmal dargestellt. Auch hier zeigt sich ein ähnliches Beziehungsmuster wie bei den beiden vorangegangenen Modellen. Ein Unterschied ist, dass in diesem Modell sowohl Lerninteresse als auch kognitive Grundfähigkeit das Lernverhalten beeinflussen und dieses nun auch vom Sozialstatus abhängt, nicht hingegen von der Erstsprache und dem Geschlecht.

³ Perfekter Modell-Fit. Erklärte Varianz auf Individualebene: $R^2 = .24$ für Leistung, $R^2 = .01$ für Hausaufgabenzeit; auf Klassenebene: $R^2 = .92$ für Leistung, $R^2 = .70$ für Hausaufgabenzeit.

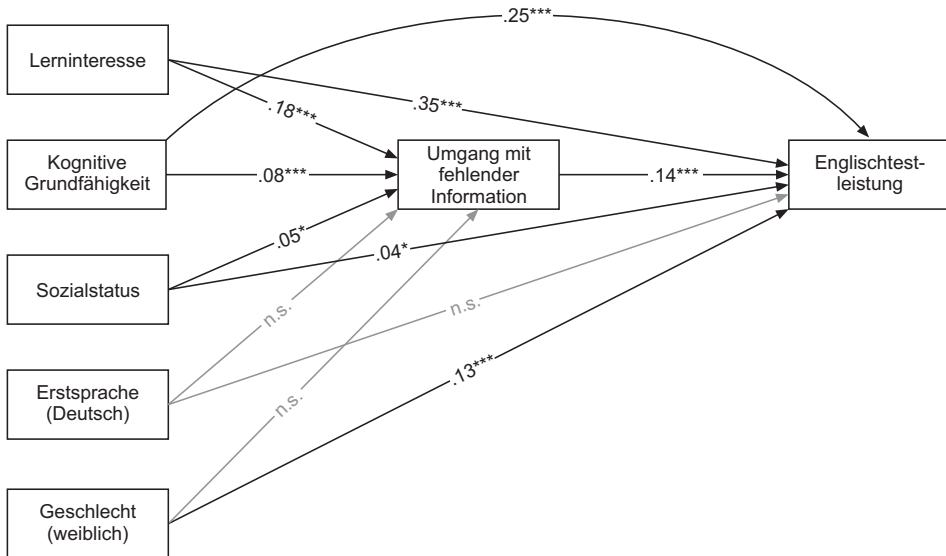


Abbildung 24.3: Pfadmodell für den Einfluss des Umgangs mit fehlender Information (SILL) auf die Englischtestleistung (Individualebene) unter Berücksichtigung von Hintergrundmerkmalen.⁴

Auf Klassenebene (ohne Abbildung) ist lediglich der Einfluss von kognitiver Grundfähigkeit, Sozialstatus und Mädchenanteil auf die Leistung signifikant. Das Klassenniveau des Lernverhaltens hängt dagegen nur mit der Sprachzusammensetzung der Klasse zusammen: Je niedriger der Anteil von Schülerinnen und Schülern mit deutscher Sprachherkunft, umso höher der Einsatz von Strategien zum Umgang mit fehlender Information.

Der Einfluss des motivationalen Merkmals ist in diesen Fällen jeweils noch etwas größer, während der von kognitiver Grundfähigkeit und Lernstrategie etwas geringer ist. Für Leseaktivitäten und Mediennutzung ergaben sich keine schätzbaren Modelle. Auch theoretisch nahe liegende Modelle mit Wechselwirkungen zwischen Merkmalen der Quantität und der Qualität des Lernverhaltens erwiesen sich als nicht schätzbar.

Im nächsten Schritt wird nun dargelegt, welche Rolle die mittels der Englischnote erfasste Eingangsleistung für das Lernen und die Leistung spielt. Die Note ist trotz der bekannten Schwächen dieses Maßes ein brauchbarer Grobindikator für den schulischen Leistungsstand; als vorteilhaft erweist sich dabei auch, dass aufgrund der moderaten Korrelation zwischen Note und Posttestleistung noch genügend Spielraum für andere Einflussgrößen bleibt.

4 Perfekter Modell-Fit. Erklärte Varianz auf Individualebene: $R^2 = .24$ für Leistung, $R^2 = .04$ für Umgang mit fehlender Information; auf Klassenebene: $R^2 = .93$ für Leistung, $R^2 = .80$ für Umgang mit fehlender Information.

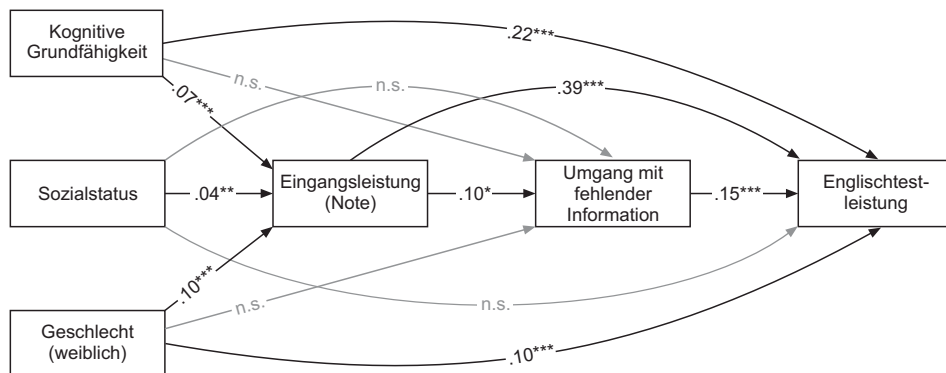


Abbildung 24.4: Pfadmodell für den Einfluss des Umgangs mit fehlender Information (SILL), Eingangsleistung (Note, umgepolt) auf die Englischtestleistung (Individual-ebene) unter Berücksichtigung von Hintergrundmerkmalen.⁵

Das Zusammenhangsmuster ist – wieder für den Umgang mit fehlender Information – in Abbildung 24.4 dargestellt. Es zeigt sich, dass die Eingangsleistung sowohl die Testleistung am Ende als auch die Lernstrategien beeinflusst. Der Einfluss von kognitiver Grundfähigkeit, Sozialstatus und Geschlecht auf die Lernstrategien verschwindet, wenn man die Eingangsleistung kontrolliert. Ob und in welchem Umfang diese Lernstrategien eingesetzt werden, hängt also weniger von der kognitiven Fähigkeit als vom Vorwissen ab.

24.3 Lernstrategien und andere lern- und leistungsrelevante Merkmale

Weitere Aufschlüsse zur Bedeutung der Lernstrategien lassen sich gewinnen, wenn man von einer variablen- zu einer personenorientierten Betrachtungsweise übergeht und unterschiedlich kompetente Lerner vergleicht. Zu diesem Zweck wurden die Schülerinnen und Schüler, getrennt nach Bildungsgang, für jedes der sechs Lernstrategiemerkmale (die drei allgemeinen Lern- und metakognitiven Strategien und die drei spezifischen Lernstrategien) in zwei Gruppen (über- und unterdurchschnittliche Ausprägung) eingeteilt. Die Schülerinnen und Schüler werden in drei unterschiedliche Lernkompetenzniveaus eingeteilt. Diejenigen, die bei allen sechs Strategien zur oberen Gruppe gehören, werden zusammengefasst, analog wird mit der unteren Gruppe verfahren; diese beiden Gruppen werden den verbleibenden Schülerinnen und Schülern gegenübergestellt. Diese drei Gruppen werden dann im Hinblick auf verschiedene lern- und leistungsrelevante Merkmale verglichen. Die für die vier Bildungsgänge resultierenden Profile sind in Abbildung 24.5 dargestellt.

5 Perfekter Modell-Fit. Erklärte Varianz auf Individualebene: $R^2 = .28$ für Testleistung, $R^2 = .02$ für Umgang mit fehlender Information, $R^2 = .02$ für Eingangsnote; auf Klassenebene: $R^2 = .93$ für Leistung, $R^2 = .75$ für Umgang mit fehlender Information, $R^2 = .25$ für die Eingangsnote.

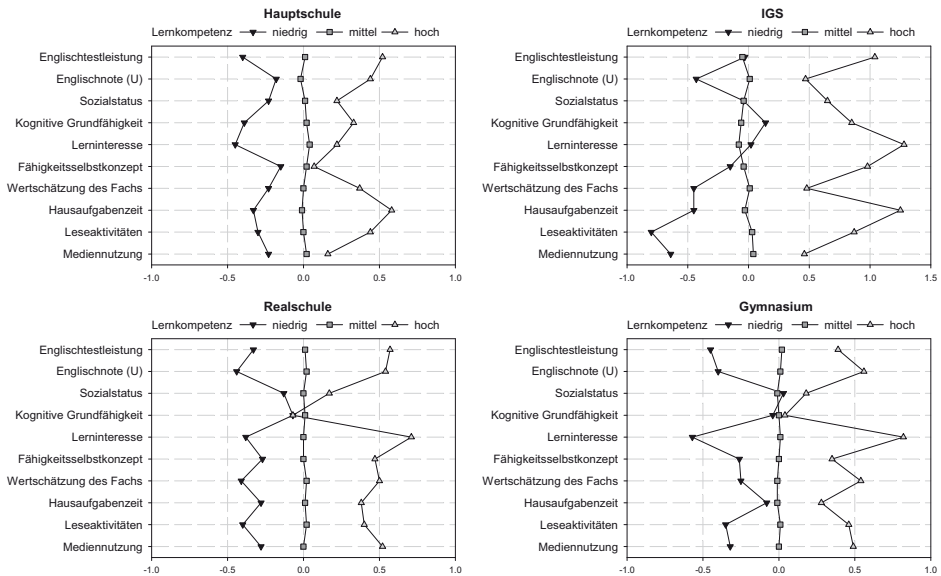


Abbildung 24.5: Merkmalsprofile für Schüler mit niedriger, mittlerer und hoher Lernkompetenz, separat nach Bildungsgang (Profilmerkmale klassenweise standardisiert).

Für alle vier Bildungsgänge zeigt sich, dass kompetente Lerner bei fast allen Merkmalen günstigere Ausprägungen aufweisen als Schüler mit mittlerer und niedriger Lernkompetenz. Aus Abbildung 24.5 ist auch ersichtlich, dass sich Schüler mit unterschiedlichen Lernstrategien in ihren Leseaktivitäten unterscheiden. Dies wirft die Frage nach möglichen Kausalbeziehungen zwischen diesen beiden Merkmalen auf. Da sowohl Lernstrategien und Leseaktivitäten jeweils zu Beginn und zu Ende der neunten Jahrgangsstufe erfasst wurden, kann dieser Frage durch die Analyse zeitverzögerter Effekte in einem cross-lagged-panel-design ansatzweise nachgegangen werden.

Abbildung 24.6 verdeutlicht, dass die Lernkompetenz (hier: der Umgang mit fehlender Information, im Wesentlichen das Erschließen und Umschreiben unbekannter englischer Wörter) im Laufe der neunten Jahrgangsstufe mäßig und die Leseaktivität (die Häufigkeit, mit der verschiedene englische Texte zum Vergnügen gelesen werden) relativ stabil bleibt. Dass sich für den Zusammenhang zwischen Lernkompetenz und Leseaktivität am Ende ein geringerer Wert ergibt als zu Beginn, liegt daran, dass es sich zum zweiten Zeitpunkt um eine partielle Korrelation handelt; die einfachen Korrelationen sind zu beiden Zeitpunkten etwa gleich hoch. Das wesentliche Ergebnis ist aber, dass die Leseaktivität zu Beginn der neunten Jahrgangsstufe einen deutlichen Einfluss auf die Lernstrategien ausübt, während die umgekehrte Einflussrichtung klar schwächer ist. Mit anderen Worten: Wer viel in der englischen Sprache liest, verbessert im Laufe der neunten Jahrgangsstufe seine Fähigkeit, unbekannte englische Wörter zu erschließen und zu umschreiben. Ist diese Fähigkeit

bereits zu Beginn der neunten Jahrgangsstufe günstig ausgeprägt, hat das allerdings kaum einen Einfluss auf die Leseaktivität am Ende der neunten Jahrgangsstufe.

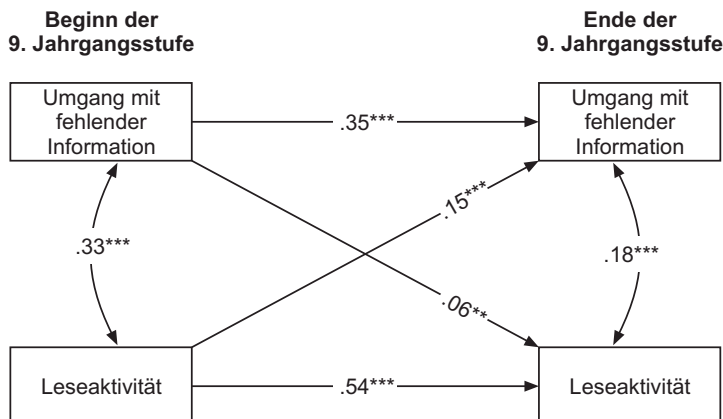


Abbildung 24.6: Beziehungen zwischen Lernkompetenz (hier: Umgang mit fehlender Information) und Leseaktivität zu Beginn und am Ende der neunten Jahrgangsstufe im Englischen.

24.4 Diskussion

Die Ergebnisse zeigen, dass im Fach Englisch spezifische Lernstrategien sowie metakognitive Strategien substantiell zur Erklärung von Leistungsunterschieden beitragen. Der Erklärungsbeitrag der allgemeinen Lernstrategien (Elaboration, Kontrolle) ist demgegenüber deutlich schwächer. Ob andere Erhebungsmethoden (z.B. die Erfassung des Lernverhaltens in der Lernsituation mittels lauten Denkens, vgl. etwa Zimmermann/Plessner 1998) zu anderen Ergebnissen führen, muss an dieser Stelle offen bleiben. Unklar ist auch, ob nicht überhaupt andere Lernstrategien als die hier berücksichtigten für die Leistung wichtiger sind. Dafür sprechen Untersuchungen mittels qualitativer Forschungsansätze (Haudeck 2007, in Vorb.).

Die in den Strukturgleichungsmodellen gefundenen direkten Pfade auf die Leistung belegen, dass Lernstrategien ein eigenständiges Erklärungsgewicht haben. Sie vermitteln darüber hinaus den Einfluss von allgemeinen Personenmerkmalen auf die Leistung. Der Effekt der Lernstrategien auf die Leistung ist allerdings geringer als der von kognitiver Grundfähigkeit und Lerninteresse, die beide gewichtige Leistungsprädiktoren darstellen. Die direkten Pfade dieser beiden Personenmerkmale machen deutlich, dass es noch andere, mittels der zugrunde liegenden Daten nicht aufklärbare Vermittlungsmechanismen gibt. Die genauere Aufhellung derartiger Prozesse muss weiterführenden Untersuchungen vorbehalten bleiben.

Die dargestellten Modelle zeigen ferner, dass kognitive Grundfähigkeit und Lerninteresse unterschiedliche Effekte auf die Anwendung von Lernstrategien haben. Beim Umgang mit fehlender Information kommen beide Merkmale zum Tragen. Für die Hausaufgabenzeit spielt nur das Lerninteresse eine Rolle, für die metakognitiven

Strategien nur die kognitive Grundfähigkeit. Motivation ist nicht nur für den Umfang des Lernverhaltens – die investierte Lernzeit –, sondern auch für dessen Qualität – die strategische Bewältigung von Lernanforderungen – relevant. Mit dem Einsatz von Lernstrategien ist eine erhöhte mentale Anstrengung verbunden, wofür ein ausreichendes Maß an Lerninteresse erforderlich ist.

Dass effektives Lernen von kognitiven Fähigkeiten abhängt, scheinen die Ergebnisse der Analysen nahezulegen. Eine detailliertere Aufschlüsselung deutet aber darauf hin, dass es möglicherweise eher das Vorwissen ist als die kognitive Fähigkeit im engeren Sinne. Im Vergleich dazu ist der Einfluss des Sozialstatus auf Individualebene – und nur diese steht hier im Blickpunkt – von untergeordneter Bedeutung. Geschlechtsunterschiede spielen dagegen eine deutlich erkennbare Rolle. Sie wirken in der erwarteten Richtung: Mädchen erzielen nicht nur bessere Englischleistungen, sondern zeichnen sich (zumindest in zwei der drei Analysen) auch durch günstigeres Lernverhalten aus.

Hausaufgabenzeiten haben wie andere Lernzeitmaße einen Doppelcharakter: Da sie von Fähigkeit und Lerninteresse in gegenläufiger Weise beeinflusst werden (je größer die Fähigkeit, umso weniger Lernzeit ist nötig, und je stärker die Motivation, umso mehr Lernzeit wird investiert), sind einfache Zusammenhänge mit der Leistung nur schwer zu interpretieren. Hier sind komplexere Modelle unabdingbar, um zu einem klareren Verständnis der Zusammenhänge zu gelangen. In den vorliegenden Analysen zeigt sich, dass die investierte Hausaufgabenzeit zwar von der Einstellung zum Fach, nicht aber von der kognitiven Fähigkeit abhängt.

In weiteren Analysen wurde versucht, den kompetenten Lerner, der durch eine überdurchschnittliche Ausprägung aller Lernkompetenzmerkmale gekennzeichnet ist, näher zu charakterisieren. Die Ergebnisse zeigen, dass sich kompetente Lerner in allen vier Bildungsgängen deutlich von weniger kompetenten unterscheiden. Deutliche Unterschiede zeigen sich vor allem bei Lerninteresse und Leseaktivitäten. Detaillierte Analysen zum Zusammenhangsgefüge dieser Merkmale müssen späteren Arbeiten vorbehalten bleiben.

Literatur

- Bimmel, P. (2006): Lernstrategien: Pläne (mental Handelns). In: Jung, U.O.H. (Hrsg.): *Praktische Handreichung für Fremdsprachenlehrer*. Frankfurt/Main: Peter Lang, S. 362-369.
- Finkbeiner, C. (1995): Zur Erhebung von textverstehensrelevanten Lernstrategien und Interessen im Fremdsprachenunterricht. Entwicklung zweier Fragebögen. In: *Empirische Pädagogik* 8, H. 2, S. 193-219.
- Haudeck, H. (1998): Lernstrategien und Lerntechniken für Schüler. In: Timm, J.-P. (Hrsg.): *Englisch lernen und lehren. Didaktik des Englischunterrichts*. Berlin: Cornelsen Verlag, S. 342-352.
- Haudeck, H. (2007, in Vorb.): *Vokabellernstrategien in der häuslichen Wortschatzarbeit im Englischen*. Dissertation: Universität Dortmund.
- Muthén, L.K./Muthén, B.O. (2004). *Mplus User's Guide. Statistical Analysis With Latent Variables*. Los Angeles, CA: Muthén & Muthén.
- Nold, G. (1992): *Lernbedingungen und Lernstrategien*. Tübingen: Gunther Narr Verlag.

- Nold, G. (2000): Kognitive, affektive und soziale Variablen beim schulischen Fremdsprachenlernen. In: Duit, R./von Rhöneck, C. (Hrsg.): *Ergebnisse fachdidaktischer und psychologischer Lehr-Lern-Forschung*, Kiel: IPN 169, S. 157-190.
- Nold, G. (2003): The impact of the learning culture on achievement in the English as a foreign language classroom – a view from Germany. In: Rymarczyk, J./Haudeck, H. (Hrsg.): *In Search of the Active Learner*. Frankfurt: Peter Lang, S. 163-184.
- Nold, G./Haudeck, H./Schnaitmann G.W. (1997): Die Rolle von Lernstrategien im Fremdsprachenunterricht. In: *Zeitschrift für Fremdsprachenforschung* 8, H. 1, S. 27-50.
- Nold, G./Schnaitmann, G.W. (1995): Lernbedingungen und Lernstrategien in verschiedenen Tätigkeitsbereichen des Fremdsprachenunterrichts. In: *Empirische Pädagogik* 9, H. 2, S. 239-261.
- Oxford, R./Burry-Stock, J. (1995): Assessing the use of language learning strategies worldwide using the Strategy Inventory for Language Learning (SILL). In: *System* 2, S. 1-23.
- Raupach, M. (2006): Müssen Schüler das Lernen lernen? In: Jung, U.O.H. (Hrsg.): *Praktische Handreichung für Fremdsprachenlehrer*. Frankfurt/Main: Peter Lang, S. 356-362.
- Schneider, W./Schlagmüller, M. (2002): Metakognitives Wissen über Textverarbeitung. Lehrstuhl für Psychologie IV, Universität Würzburg.
- Schrader, F.-W./Helmke, A./Hosenfeld, I./Wagner, W. (2003, März): Allgemeine oder bereichsspezifische Lernstrategien? Untersuchungen zur Validität des Strategy Inventory for Language Learning (SILL). Forschungsreferat bei der Frühjahrstagung der Arbeitsgemeinschaft für empirische pädagogische Forschung (AEPF) in Frankfurt/Main.
- Trautwein, U./Köller, O./Schmitz, B./Baumert, J. (2002): Do homework assignments enhance achievement? A Multilevel analysis in 7th-grade mathematics. In: *Contemporary Educational Psychology* 27, S. 26-50.
- Wolff, D. (1998): Lernerstrategien beim Fremdsprachenlernen. In: Timm, J.-P. (Hrsg.): *Englisch lernen und lehren. Didaktik des Englischunterrichts*. Berlin: Cornelsen Verlag, S. 70-77.
- Zimmermann, G./Plessner, H. (1998): Zum Zusammenhang von Lernstrategien, Wissensrepräsentationen und Leistungen beim Lernen mit Instruktionstexten im Fremdsprachenunterricht. In: *Zeitschrift für Fremdsprachenforschung* 9, H. 2, S. 265-290.

Hans-Günter Rolff / Michael Leucht / Ernst Rösner

25 Sozialer und familialer Hintergrund

25.1 Einführung

Die DESI-Studie ist nicht die erste Untersuchung, die einen Zusammenhang zwischen den in ihr erhobenen Daten und den sozialen wie familialen Hintergrundbedingungen der betrachteten Schülerschaft herzustellen versucht. Die Bildungssoziologie hat diesen Zusammenhang bereits in den 60er Jahren untersucht (vgl. Coleman 1966; Rolff 1967). Studien wie PISA (vgl. Baumert u.a. 2001; Prenzel u.a. 2004) haben dies unter Einbezug internationaler Stichproben geleistet und vertieft. Auch in auf den bundesdeutschen Raum bezogenen Studien (z.B. Köller u.a. 2004) bildet in den letzten Jahren die soziale Herkunft von Schülerinnen und Schülern einen prominenten Untersuchungsbereich.

„Klassische“ Untersuchungsbereiche – Disparitäten in Bildungsbeteiligung und Kompetenzerwerb

Mit sozialer Herkunft werden in diesen Untersuchungen zumindest zwei Sachverhalte in Verbindung gebracht: Bildungsbeteiligung und Kompetenzen von Schülerinnen und Schülern. Wird der Zusammenhang zwischen *Bildungsbeteiligung* und sozialer Herkunft betrachtet, ist etwa die Frage gestellt, ob sich Schüler aus Familien mit unterschiedlichen sozialen und familialen Hintergrundbedingungen im Hinblick auf ihre Zugangschancen zu bestimmten Bereichen des Bildungssystems unterscheiden. Geht es hingegen um den Zusammenhang von *Kompetenzen* und sozialer Herkunft, würde etwa gefragt, ob sich Schüler aus verschiedenen sozialen Gefügen im Hinblick auf ihre Leistungen in einzelnen Schulfächern unterscheiden. Ungleichheiten in Bezug auf Bildungsbeteiligung (vgl. etwa Baumert/Schümer 2001) oder erworbene Kompetenzen (siehe zuletzt Ehmke/Siegle/Hohensee 2005) werden als *Disparitäten* bezeichnet.

Theoretisches Rahmenmodell sozialer Herkunft – Die Kapitalterminologie Bourdieus

Empirisch ermittelten Disparitäten steht auf theoretischer Ebene ein bildungssoziologisches Erklärungsmodell gegenüber, dass sich auf Arbeiten Bourdieus (vgl. z.B. 1983) sowie Colemans (z.B. 1988) stützt. Im Hinblick auf Hintergrundbedingungen sozialer Herkunft unterscheiden diese drei Arten von Ressourcen: Ökonomisches, soziales und kulturelles Kapital. Das *ökonomische Kapital* der Familie bezeichnet neben dem Besitz von Produktionsmitteln, Grund und Boden auch Geldvermögen, Aktien und alle anderen unmittelbar in Geld konvertierbaren Gegenstände. *Soziales Kapital* meint Ressourcen, die sich aus dem mehr oder weniger engen Netz sozialer

und familialer Beziehungen ergeben. Soziales Kapital wird vornehmlich in Familien, im Freundeskreis, in Vereinen oder anderen Gruppen erworben. *Kulturelles Kapital* schließlich fasst Ressourcen, die die Grundlage für eine regelmäßige Teilhabe an der bürgerlichen Kultur darstellen. Im Hinblick auf das kulturelle Kapital unterscheidet Bourdieu objektivierte, institutionalisiertes und verinnerlichtes bzw. inkorporiertes kulturelles Kapital. Während objektiviertes kulturelles Kapital etwa im Besitz an Kulturgütern und institutionalisiertes kulturelles Kapital in den innerhalb der Familie erreichten Bildungsabschlüssen zum Ausdruck kommt, äußert sich inkorporiertes kulturelles Kapital in Wahrnehmungs-, Denk- und Handlungsschemata von Personen innerhalb der Familie¹.

Die drei Kapitalarten sind nicht unabhängig voneinander. So ist ökonomisches Kapital eine der Voraussetzungen zur Bildung der anderen beiden Kapitalarten. Andererseits sind soziales und kulturelles Kapital in ökonomisches Kapital konvertierbar (siehe Portes 1998).

Erfassung kapitalterminologisch gefasster sozialer Herkunft

Zur Erfassung kapitalterminologisch gefasster sozialer Herkunft werden üblicherweise mehrere Indikatoren herangezogen. So wird ökonomisches Kapital über die sozioökonomische Stellung der Familie und den Erwerbstätigkeitsstatus des Hauptverdieners, soziales Kapital über Indikatoren der Familienstruktur und kulturelles Kapital über die von Eltern erreichten Bildungsabschlüsse oder den Besitz an Kulturgütern (z.B. Bücher, Musikinstrumente) operationalisiert². So wurde auch in DESI vorgegangen.

Struktur- und Prozessmerkmale sozialer Herkunft

In neueren Untersuchungen (vgl. z.B. Baumert/Watermann/Schümer 2003) hat es sich eingebürgert, Struktur- und Prozessmerkmale sozialer Herkunft zu unterscheiden. Als Strukturmerkmale werden zeitstabile Hintergrundbedingungen wie der sozioökonomische Status, das Bildungsniveau von Eltern sowie der Migrationsstatus genannt. Prozessmerkmale hingegen sind solche, deren Ausprägung von diesen Strukturmerkmalen abhängig ist – also etwa Werthaltungen und Einstellungen der Eltern oder konkrete Formen elterlicher Unterstützung (siehe etwa Helmke/Schrader/Hosenfeld 2004). Hier, also auf Prozessebene, würden diejenigen Ressourcen erzeugt, die erfolgreiche Bildungsverläufe der jüngeren Generation unterfüttern. Entsprechend wird versucht (Baumert u.a. 2003), strukturelle und prozessuale Aspekte vor allem kulturellen Kapitals separat erfassbar zu machen. Grundlegende Annahme ist dabei,

1 Vgl. hierzu auch Bourdieus Konzept des *Habitus*.

2 In letzter Zeit (vgl. Ehmke u.a. 2004) wurde mit dem so genannten ESCS (Index of Economical and Socio-Cultural Status) ein Indikator vorgelegt, der eine simultane Erfassung der sozioökonomischen Stellung, elterlicher Bildungsabschlüsse sowie kultureller Besitztümer in der Familie erlaubt.

dass die Effekte struktureller Bedingungen sozialer Herkunft auf Schülerleistung zumindest zum Teil durch Prozessmerkmale vermittelt werden.

Neue Indikatoren zur Erfassung von Prozessmerkmalen sozialer Herkunft

Zur Erfassung von Prozessmerkmalen sozialer Herkunft wurden in DESI zwei Arten von Indikatoren herangezogen. Sie erfragen einerseits auf die Sprachen Deutsch und Englisch bezogene Werthaltungen und Einstellungen der Eltern (*Sprachpraxis im Elternhaus*). Entsprechende Fragen bezogen sich etwa auf Deutsch- und Englischkenntnisse in der Familie (z.B. „Ich glaube, in Deutsch / Englisch wissen wir so gut Bescheid wie die Lehrer unseres Kindes.“), auf den innerfamiliären Umgang mit diesen Sprachen (z.B. für Deutsch: „In unserer Familie gibt es Gespräche über sprachliche Fragen, zum Beispiel: ‚Kann man das so sagen oder schreiben?‘“) sowie Elterneinstellungen zu sprachbezogenen Unterstützungsformen („Wenn unser Kind in Deutsch / Englisch nicht mehr weiter weiß, helfen wir ihm.“). Andererseits wurden die *tatsächlichen sprachbezogenen Unterstützungsformen* (z.B. Vokabelabfrage in Englisch) in ihrer jeweiligen Häufigkeit (Antwortvorgaben: „nie“, „mehrmals im Monat“, „mehrmals pro Woche“, „täglich“) erfasst.

Derartige sprachbezogene Einstellungen sowie tatsächliche sprachbezogene Unterstützungsformen wurden von Bourdieu und Coleman so gut wie nicht thematisiert. Dennoch ist im Anschluss an Bernstein (1977) zu vermuten, dass sie durch Strukturmerkmale sozialer Herkunft geprägt werden und als prozessuale Aspekte kulturellen Kapitals Einfluss auf sprachliche Kompetenzen von Schülerinnen und Schülern haben. Unter Nutzung der obigen Unterscheidung (vgl. Bourdieu 1983) wären Sprachpraxis im Elternhaus und tatsächliche sprachbezogene Unterstützungsformen als inkorporiertes kulturelles Kapital aufzufassen, also als verinnerlichte Wahrnehmungs-, Denk- und Handlungsschemata von Personen in der Familie.

25.2 Überblick über das Kapitel

Im folgenden Kapitel geht es einerseits um die Replikation der auf strukturelle Aspekte sozialer Herkunft bezogenen Analysen aus PISA 2003 (vgl. Ehmke u.a. 2004; Ehmke/Siegle/Hohensee 2005), andererseits um deren Arrondierung durch die in DESI erstmals betrachteten sprachbezogenen Prozessmerkmale sozialer Herkunft.

Deskription von Strukturmerkmalen der sozialen Herkunft

Eine Deskription struktureller Aspekte sozialer Herkunft soll die Frage beantworten, inwieweit in der in DESI betrachteten Schülerschaft im Hinblick auf Familienstruktur, sozioökonomische Stellung der Familie, Erwerbstätigkeit der

Eltern, elterliche Bildungsabschlüsse sowie kulturelle Besitztümer innerhalb der Familie mit PISA 2003 vergleichbare Bedingungen vorliegen. Eine Übereinstimmung der Bedingungsgefüge ist Voraussetzung für einen sinnvollen Vergleich der Effekte dieser Gefüge auf Schülerkompetenzen in Mathematik (PISA 2003) und sprachlichen Fächern (DESI).

Deskription sprachbezogener Prozessmerkmale sozialer Herkunft

Durch eine Deskription der Sprachpraxis im Elternhaus und der tatsächlichen sprachbezogenen Unterstützungsformen sollen hier erstmals sprachbezogene Prozessmerkmale der sozialen Herkunft beleuchtet werden. Zudem ist durch die Beschreibung auf die Fächer Deutsch und Englisch bezogener Elterneinstellungen und Unterstützungsformen in Familien aus unterschiedlichen strukturellen Bedingungen sozialer Herkunft intendiert, die inhaltliche Plausibilität der neu gebildeten Indikatoren zu illustrieren.

Mittlere Sprachleistungen von Schülern aus verschiedenen sozialen Bedingungen

Durch Replikation der Analysen von Ehmke und anderen (2004) soll der Frage nachgegangen werden, welche mittleren sprachlichen Leistungen Schüler aufweisen, die zuvor bestimmten Bedingungen sozialer Herkunft zugeordnet wurden. Dies ermöglicht die Gesamtschau entsprechender Befunde für die drei Hauptfächer Deutsch, Englisch und Mathematik. Der Zusammenhang von sozioökonomischem Status der Familie und Kompetenzen in Deutsch und Englisch wird, wiederum unter Bezugnahme auf vergleichbare Ergebnisse aus PISA 2003, zudem durch Analysen zum sozialen Gradienten³ näher untersucht.

Struktur- und Prozessmerkmale sozialer Herkunft und Schülerleistungen in sprachlichen Fächern – ein Mediationsmodell

Schließlich werden Schülerleistungen in Deutsch und Englisch zu Strukturmerkmalen und Prozessmerkmalen der sozialen Herkunft regressionsanalytisch in Beziehung gesetzt. Diese Analysen dienen erstens – erneut mit Rückblick auf PISA 2003 – einer Gesamtschau von Effekten struktureller Aspekte sozialer Herkunft auf Schülerleistungen in den Fächern Deutsch, Englisch und Mathematik. Zweitens soll durch sie geklärt werden, inwieweit die Sprachpraxis im Elternhaus als Bündel von Prozessmerkmalen sozialer Herkunft auf Sprachleistungen einwirkt und den Einfluss struktureller Aspekte sozialer Herkunft mediiert.

3 Bei PISA 2000 und 2003 wurde zur Quantifizierung des Zusammenhangs zwischen sozialer Herkunft der Familie und Schülerkompetenz der so genannte soziale Gradient herangezogen. Er bringt in Form einer einfachen Regression die Kopplung des höchsten sozioökonomischen Status der Familie (HISEI; PISA 2000) bzw. ESCS (PISA 2003) mit Fachleistungen zum Ausdruck.

25.3 Deskription von Strukturmerkmalen der sozialen Herkunft

Zur Deskription struktureller Aspekte der sozialen Herkunft werden die folgenden Indikatoren herangezogen:

- Erwerbstätigkeit von Mutter und Vater im Sinne des Anteils von Vollzeit- und Teilzeitbeschäftigten bzw. Arbeitssuchenden,
- Familienstruktur, wobei Anteile so genannter Kernfamilien und allein erziehender Elternhäuser ausgewiesen sind,
- Höchster sozioökonomischer Status in der Familie (HISEI; vgl. Ganzeboom u.a. 1992),
- Höchster elterlicher Bildungsabschluss gemäß der International Standard Classification of Education (ISCED; vgl. OECD 1999),
- Ein integrierter Indikator zum kulturellen Besitz in der Familie, der Angaben zum Besitz von Kulturgütern (z.B. klassische Literatur, Kunstwerke) und zum Buchbestand in der Familie bündelt und einen theoretischen Wertebereich von Null bis 1 aufweist.

Zur besseren Vergleichbarkeit wurden alle Indikatoren wie in PISA 2003 (Ehmke u.a. 2004) operationalisiert. Tabelle 25.1 zeigt Ausprägungen der strukturellen Aspekte der sozialen Herkunft für die in DESI betrachtete Schülerschaft sowie aufgefächert für Schüler aus den vier in DESI betrachteten Bildungsgängen. Betreffs eines Vergleiches der in DESI vorgefundenen Ausprägungen struktureller Aspekte sozialer Herkunft mit PISA 2003 (vgl. Ehmke u.a. 2004) kann Tabelle 25.1 entnommen werden: Im Hinblick auf die Erwerbstätigkeit von Mutter und Vater, die Familienstruktur, den höchsten sozioökonomischen Status der Familie, höchste elterliche Bildungsabschlüsse und den Besitz von Kulturgütern wurde eine in hohem Maße mit PISA 2003 vergleichbare Konstellation ermittelt. Statistisch bedeutsame Abweichungen ergeben sich lediglich für den Anteil teilzeitbeschäftigter Väter, der in DESI etwas höher ausfällt, sowie für den höchsten elterlichen Bildungsabschluss gemäß ISCED, dessen niedrigste Kategorie in DESI etwas stärker besetzt ist.

Im Hinblick auf die Ausprägungen der betrachteten Indikatoren in den vier Bildungsgängen Haupt-, Real-, Gesamtschule und Gymnasium kann – in Übereinstimmung mit Befunden aus PISA 2000 – festgehalten werden, dass Familien von Haupt- und Gesamtschülern insgesamt ungünstigere, Familien von Realschülern und Gymnasiasten dagegen positivere Bedingungsgefüge aufweisen.

Tabelle 25.1: Ausprägungen von Indikatoren der sozialen Herkunft in der DESI-Stichprobe und den vier betrachteten Bildungsgängen.

Indikatoren sozialer Herkunft	Mittelwerte und Standardfehler nach Bildungsgängen				
	Gesamt	HS	RS	GS	GYM
Erwerbstätigkeit Vater^A					
Vollzeit	80.4 (0.5)	75.9 (1.4)	80.6 (0.8)	75.1 (1.7)	85.6 (0.9)
Teilzeit	11.5 (0.5)	13.5 (1.3)	11.8 (0.8)	15.4 (1.5)	8.3 (0.7)
Arbeitssuchende	4.3 (0.3)	5.0 (0.6)	4.8 (0.5)	5.3 (0.9)	2.7 (0.4)
Erwerbstätigkeit Mutter^A					
Vollzeit	32.4 (0.8)	29.5 (1.4)	32.4 (1.6)	38.3 (4.0)	33.0 (1.6)
Teilzeit	43.4 (0.9)	41.2 (2.1)	44.1 (1.5)	37.1 (3.1)	46.1 (1.6)
Arbeitssuchende	5.3 (0.4)	6.6 (1.1)	5.6 (0.6)	6.8 (1.2)	3.3 (0.5)
Familienstruktur^B					
Alleinerziehende	17.6 (0.7)	23.4 (2.0)	15.2 (0.8)	24.0 (2.5)	13.6 (0.8)
Kernfamilien	73.7 (0.7)	66.1 (2.0)	75.8 (0.9)	67.0 (2.9)	79.8 (1.0)
Sozioökonom. Status (HISEI)	48.4 (0.3)	41.2 (0.4)	47.3 (0.4)	44.9 (1.0)	57.4 (0.4)
Bildungsabschluss (HISCED)					
Primar/Sekundarstufe I	20.5 (0.6)	34.5 (1.6)	20.1 (0.9)	24.4 (3.1)	8.9 (0.6)
Sekundarstufe II	43.1 (0.7)	41.8 (1.2)	47.1 (1.1)	47.0 (2.3)	38.1 (1.3)
Tertiärstufe	36.4 (0.7)	23.8 (1.3)	32.8 (1.0)	28.6 (3.4)	53.0 (1.2)
Besitz von Kulturgütern	0.48 (0.0)	0.35 (0.0)	0.45 (0.0)	0.43 (0.0)	0.65 (0.0)

Anmerkungen: ^AJeweils ohne die Kategorie „Etwas anderes (Hausfrau/-mann, Rentner/in)“, ^BOhne so genannte Mischfamilien.

25.4 Deskription sprachbezogener Prozessmerkmale sozialer Herkunft

Für die Deskription sprachbezogener Prozessmerkmale sozialer Herkunft werden zunächst auf die Sprachen Deutsch und Englisch bezogene Werthaltungen und Einstellungen von Eltern (Sprachpraxis im Elternhaus) herangezogen. Daran anschließend werden verschiedene Formen tatsächlicher sprachbezogener Unterstützung innerhalb der Familie betrachtet.

Deskription sprachbezogener Werthaltungen und Einstellungen von Eltern (Sprachpraxis im Elternhaus)

Zur Erfassung auf die Sprachen Deutsch und Englisch bezogener Werthaltungen und Einstellungen von Eltern wurden die folgenden sieben Skalen herangezogen:

- Kenntnisse in der Familie (Beispielitem: „Ich glaube, in Deutsch / Englisch wissen wir so gut Bescheid wie die Lehrer unseres Kindes.“),
- Bedeutung von Deutsch / Englisch im Beruf der Mutter (Beispielitem: „Ich habe in meinem Beruf viel mit deutschen bzw. englischen Texten zu tun.“),

- Bedeutung von Deutsch / Englisch im Beruf des Vaters
- Unterstützung durch die Familie (Beispielitem: „Wenn unser Kind in Deutsch / Englisch nicht mehr weiter weiß, helfen wir ihm.“),
- Umgang mit Deutsch / Englisch in der Familie (Beispielitems: „In unserer Familie gibt es Gespräche über sprachliche Fragen, z.B. „Kann man das so sagen oder schreiben“. bzw. „Wir schauen uns gemeinsam amerikanische oder englische Filme in Originalsprache an.“),
- Wertschätzung von Deutsch und Englisch (Beispielitem: „In unserer Familie betrachten wir Deutsch / Englisch als wichtiges Fach.“),
- Elternteress an Deutsch- / Englischunterricht (Beispielitem: „Wir fragen unser Kind, was es im Deutschunterricht gemacht hat.“).

Tabelle 25.2 zeigt Mittelwerte und Standardabweichungen der auf die Fächer Deutsch und Englisch bezogenen Sprachpraxis im Elternhaus in der in DESI betrachteten Gesamtelternschaft⁴.

Tabelle 25.2: Ausprägungen auf die Fächer Deutsch und Englisch bezogener Sprachpraxis im Elternhaus (Mittelwerte und Standardabweichungen).

Elterneinschätzungen zu ...	Deutsch	Englisch
... sprachbezogenen Kenntnissen der Familie	2.9 (0.6)	2.1 (0.7)
... Bedeutung der Sprache im Beruf der Mutter	3.3 (0.8)	1.5 (0.7)
... Bedeutung der Sprache im Beruf des Vaters	3.2 (0.8)	1.7 (0.9)
... sprachbezogener Unterstützung in der Familie	3.1 (0.7)	2.2 (0.8)
... Umgang mit Deutsch / Englisch in der Familie	3.0 (0.7)	1.9 (0.6)
... Wertschätzung der Sprache in der Familie	3.7 (0.5)	3.6 (0.5)
... Elterninteresse am Deutsch- / Englischunterricht	3.1 (0.7)	2.9 (0.8)

Die Daten weisen aus, dass im Hinblick auf das Fach Deutsch eine bessere Sprachpraxis im Elternhaus ermittelt werden konnte als für das Fach Englisch. Dies zeigt sich in höheren innerfamiliären Kenntnissen und einem konstruktiveren Umgang als mit Englisch. Die Sprache Deutsch spielt in den Berufen von Mutter und Vater (im Sinne des Rezipierens und Produzierens von Texten) eine deutlich größere Rolle als Englisch. Für Deutsch wird ein höheres Ausmaß an innerfamiliärer Unterstützung berichtet. Wertschätzung und Elterninteresse am Unterricht beider Sprachen unterscheiden sich nicht, wobei für erstere die insgesamt höchsten Werte berichtet werden. Dieser letzte Befund wird kontrastiert durch Ergebnisse von Ehmke und anderen aus PISA 2003 (vgl. Ehmke u.a. 2004): Dort wurden für die innerfamiliäre Wertschätzung des Faches Mathematik, erfasst über Schülerangaben, sehr niedrige Werte ermittelt.

Zusammenhänge zu strukturellen Aspekten sozialer Herkunft: Wenn die Skalen zur Sprachpraxis im Elternhaus als Prozessmerkmale sozialer Herkunft aufgefasst

⁴ Die Mittelwerte weisen einen theoretischen Wertebereich von 1 (alle Items der Skala wurden mit „völlig falsch“ bzw. „nie“ beantwortet) bis 4 („völlig richtig“ bzw. „fast immer“) auf.

werden, sollten deren Ausprägungen mit denen von strukturellen Aspekten sozialer Herkunft kovariieren. Abbildung 25.1 zeigt die Mittelwerte der sieben Skalen zur Sprachpraxis im Elternhaus für das Fach Englisch in Familien, die das unterste (0-25%) bzw. das oberste Quartil (75-100%) der Verteilung des Indikators zu kulturellen Besitztümern in der Familie bilden. Es zeigt sich, dass die von Eltern berichteten sprachbezogenen Werthaltungen und Einstellungen in Familien, die viel objektiviertes kulturelles Kapital ihr Eigen nennen, durchgängig um etwa eine Standardabweichung (vgl. Tabelle 25.2) höher liegen als in Familien mit wenigen kulturellen Besitztümern. Diese Unterschiede sind hochsignifikant. Analoge Ergebnisse ergeben sich für die auf das Fach Deutsch bezogenen Skalen zur Sprachpraxis im Elternhaus. Ebenso konnten positive Zusammenhänge von auf Deutsch und Englisch bezogener Sprachpraxis im Elternhaus und höchstem sozioökonomischen Status der Familie (HISEI) bzw. höchsten elterlichen Bildungsabschlüssen (HISCED) belegt werden.

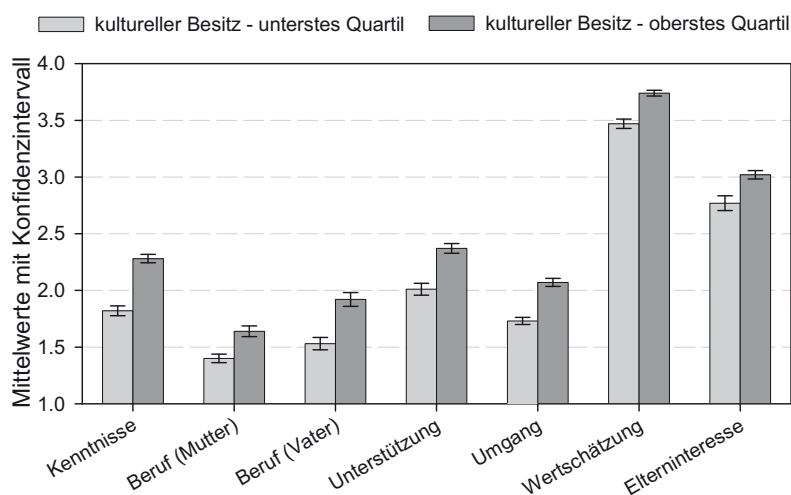


Abbildung 25.1: Auf das Fach Englisch bezogene Sprachpraxis im Elternhaus in Familien mit viel und wenig objektiviertem kulturellen Kapital.

Eltern aus Familien, die sich durch positive strukturelle Bedingungen sozialer Herkunft kennzeichnen lassen, berichten demnach auch positivere sprachbezogene Einstellungen und Werthaltungen. Umgekehrt werden bei Eltern aus Familien mit niedrigem sozioökonomischen Status, niedrigeren elterlichen Bildungsabschlüssen und wenig kulturellem Besitz geringere sprachbezogene Wertschätzungen, Selbsteinschätzungen im Hinblick auf entsprechende Kenntnisse und weniger positive sprachbezogene Umgangsformen ermittelt.

Deskription tatsächlicher sprachbezogener Unterstützung in der Familie

Wenden wir uns nun tatsächlichen sprachbezogenen Unterstützungsformen in der Familie zu. Mit dem DESI-Elternfragebogen konnte die konkrete häusliche Unterstützung der Schülerinnen und Schüler in den Fächern Deutsch und Englisch ermittelt werden. Im Einzelnen ging es um Hausaufgabenhilfe, Vokabelabfrage, Geschwisterhilfe und Nachhilfeunterricht. Bei der Auswertung der Daten wurde ergänzend der Frage nachgegangen, ob die Ausprägung häuslicher Unterstützung nach erkannter Hilfebedürftigkeit der Schülerinnen und Schüler variiert. Exemplarisch werden hier die von den Eltern angegebenen Häufigkeiten („nie“, „mehrmals im Monat“, „mehrmals pro Woche“ bzw. „täglich“) von auf die Fächer Deutsch und Englisch bezogener Hausaufgabenhilfe betrachtet.

Weder für das Fach Deutsch noch für das Fach Englisch gehört Hausaufgabenhilfe zu den verbreiteten Unterstützungsformen im Elternhaus. 65% der befragten Eltern geben an, im Fach Deutsch niemals Hausaufgabenhilfe zu leisten, im Fach Englisch sind es sogar 71%. Mehrmals pro Monat unterstützen 30% der Eltern ihre Kinder bei Deutsch-Hausaufgaben, bei Englisch-Hausaufgaben sind es 24%. Die gelegentlich geäußerte Vermutung, die Anforderungen der Schule drängten Eltern in die Rolle von Hilfslehrern, kann nach dieser Datenlage nicht bestätigt werden.

Vergleichbares gilt für die verbreitete Annahme, mit dem Anspruchsniveau des besuchten Bildungsgangs nehme das Ausmaß elterlicher Unterstützung zu. Nach den Ergebnissen von DESI verhält es sich eher umgekehrt. 64% der Eltern von Hauptschülern erteilen niemals Hausaufgabenhilfe im Fach Deutsch, bei Eltern von Realschülern sind es 64% und bei Eltern von Gymnasiasten 76%. Für Eltern von Gesamtschülern wurde die insgesamt häufigste Hausaufgabenhilfe im Fach Deutsch ermittelt (39% der Angaben entfallen auf „mehrmals im Monat“). Für die auf das Fach Englisch bezogene Hausaufgabenhilfe in den verschiedenen Bildungsgängen ergibt sich ein identisches Muster.

Erwartungsgemäß erbringt die Auswertung des Ausmaßes der Hausaufgabenhilfe nach leistungsstarken und leistungsschwachen Schülerinnen und Schülern keine grundlegend anderen Befunde. Die vorfindlichen Daten dokumentieren lediglich, dass die verbreitete Nichtbefassung mit Hausaufgabenhilfe geringfügig schwächer ausgeprägt ist, wenn die Leistung der Kinder als unbefriedigend wahrgenommen wird.

Die Ergebnisse zu den anderen Bereichen häuslicher Unterstützung sind weitgehend ähnlich. So geben mehr als die Hälfte der befragten Eltern an, niemals Vokabeln abzufragen. In der Tendenz zeigt sich das bereits beschriebene Bild, wonach die Unterstützung resp. die Intervention der Eltern umso geringer ist, je höherwertiger der besuchte Bildungsgang ihres Kindes ist. Dazu passt die Feststellung eines Zusammenhangs mit der Umgangssprache im Elternhaus: In Elternhäusern ohne Deutsch als Kommunikationssprache erfolgen Vokabelabfragen doppelt so oft wie in Elternhäusern mit Deutsch als alleiniger Sprache.

Inwieweit die Befunde zur Nutzung von Nachhilfeunterricht nach Maßgabe sozialer Erwünschtheit verzerrt sein könnten, bleibt spekulativ. Gemessen an der Verbreitung von Nachhilfeunterricht in Deutschland kann nicht ausgeschlossen werden, dass die Auskünfte der Eltern die vorfindliche Realität nicht in vollem Umfang widerspiegeln: 95% erhalten der Datenlage zufolge niemals Nachhilfeunterricht im Fach Deutsch, im Fach Englisch sind es 89%.

Zur Geschwisterhilfe teilen 80% der Eltern mit, dass keine Geschwister im Haushalt leben, die ihrem Kind beim Lernen helfen könnten. Wo es aber potenziell helfende Geschwister gibt, ist deren Unterstützung eher gering. In den Fächern Deutsch und Englisch helfen nur fünf bis sechs Prozent der Geschwister regelmäßig bei Hausaufgaben, minimal häufiger bei der Vorbereitung auf Klassenarbeiten, etwas seltener bei der Korrektur von Klassenarbeiten.

Zusammenhänge zu strukturellen Aspekten sozialer Herkunft: Inwieweit ist die (eher seltene) sprachbezogene Unterstützung abhängig von strukturellen Aspekten sozialer Herkunft? Die obigen Befunde zu sprachbezogenen Werthaltungen und Einstellungen lassen erwarten, dass Familien aus positiven strukturellen Bedingungen sozialer Herkunft auch durch ein höheres Maß tatsächlicher sprachbezogener Unterstützung gekennzeichnet werden können. Werden nun die Häufigkeiten von Hausaufgabenhilfe in Deutsch und Englisch mittels einfacher Regression durch den höchsten sozioökonomischen Status in der Familie (HISEI), den höchsten elterlichen Bildungsabschluss (HISCED) oder den kulturellen Besitztümern in der Familie vorhergesagt, können damit jeweils etwa als 0,5% der beobachteten Häufigkeitsunterschiede erklärt werden. Die statistisch dennoch signifikanten Regressionsgewichte, die sich bei diesen Rechnungen ergeben, sind negativ. Bei insgesamt nur schwachen Zusammenhängen zwischen Strukturmerkmalen sozialer Herkunft und tatsächlicher sprachbezogener Unterstützung gilt also: Je positiver die strukturellen Bedingungen sozialer Herkunft in einer Familie sind, desto seltener wird tatsächliche sprachbezogene Unterstützung, etwa im Form von Hausaufgabenhilfe, berichtet.

Zusammenhänge zu den von Eltern wahrgenommenen Leistungen ihres Kindes: Helmke, Schrader und Hosenfeld (2004) haben zeigen können, dass sich das Ausmaß lernbezogener Unterstützung durch Eltern danach unterscheidet, wie Eltern die Leistungen ihres Kindes wahrnehmen. Und tatsächlich zeigen sich auch in DESI zwischen der fachbezogenen Zufriedenheit von Eltern mit den Leistungen ihres Kindes (Beispielitem: „Wie zufrieden sind Sie mit den Leistungen ihres Kindes im Fach Deutsch bzw. Englisch?“) und den von Eltern berichteten Häufigkeiten etwa von Hausaufgabenhilfe negative Zusammenhänge ($r_{\text{Deutsch}} = -.19$; $r_{\text{Englisch}} = -.16$; beide $p < .01$). Je unzufriedener Eltern also mit den Leistungen ihres Kindes sind, desto häufiger unterstützen sie es bei seinen Hausaufgaben⁵. Dieser offensichtlich kom-

5 Der negative Zusammenhang zwischen Leistungszufriedenheit und sprachbezogenen Unterstützungsformen bleibt auch dann erhalten, wenn Strukturmerkmale sozialer Herkunft (HISEI, kulturelle Besitztümer) auspartialisiert werden.

pensatorische Charakter der erfragten Unterstützungsformen steht in Einklang mit Befunden der oben genannten Studie (Helmke/Schrader/Hosenfeld 2004).

25.5 Mittlere Sprachleistungen von Schülern aus verschiedenen sozialen Bedingungen

Hier werden mittlere Deutsch- und Englischleistungen von Schülern betrachtet, die verschiedenen strukturellen Bedingungen sozialer Herkunft zugeordnet wurden. Zur besseren Vergleichbarkeit werden zusätzlich entsprechende, auf das Fach Mathematik bezogene Ergebnisse aus PISA 2003 (vgl. Ehmke u.a. 2004) referiert⁶. Abbildung 25.2 zeigt entsprechende Werte für Schüler, deren Familien den vier Quartilen von HISEI, den Quartilen des Indikators zu kulturellen Besitztümern sowie den drei HISCED-Stufen zugeordnet wurden.

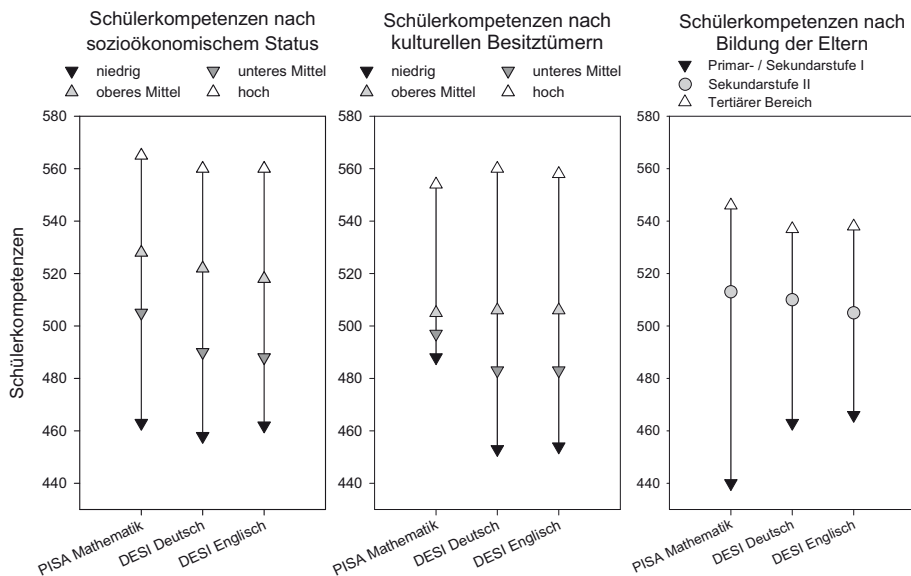


Abbildung 25.2: Mittlere Kompetenzen in Deutsch, Englisch und Mathematik (Werte aus PISA 2003) nach höchstem sozioökonomischen Status, kulturellen Besitztümern und höchstem elterlichen Bildungsabschluss.

Abbildung 25.2 veranschaulicht, dass der höchste sozioökonomische Status der Familie auf deskriptiver Ebene mit sprachlichen Leistungen in einem ebenso starken

6 Die PISA-Ergebnisse beziehen sich allerdings auf eine internationale Kompetenzskala, auf der Deutschland einen Mittelwert von 503 und eine Standardabweichung von 103 besitzt, während die beiden DESI-Gesamtskalen innerhalb Deutschlands auf den Mittelwert 500 und die Standardabweichung 100 normiert wurden.

Zusammenhang steht wie mit mathematischen Kompetenzen⁷. Nach objektiviertem kulturellen Kapital – im Sinne des Besitzes von Symbolen klassisch-bürgerlicher Bildung im Elternhaus – gruppierte Schüler zeigen in Deutsch und Englisch größere Leistungsunterschiede als im Hinblick auf mathematische Kompetenzen. Schüler aus Familien, deren Eltern unterschiedliche Bildungszertifikate erreicht haben, weisen in Bezug auf das Fach Mathematik eine breitere Leistungsstreuung auf als in den sprachlichen Fächern.

Tabelle 25.3 repliziert entsprechende Analysen für den Migrationshintergrund⁸ (für vertiefende Analysen vgl. Kapitel 20) sowie die Umgangssprache in der Familie („Umgangssprache ist Deutsch“ vs. „Umgangssprache ist nicht Deutsch“).

Tabelle 25.3: Mittlere Sprachleistungen nach Migrationshintergrund bzw. Umgangssprache in der Familie.

Schülerleistungen nach Migrationshintergrund der Familie				
	Kein Migrationshintergrund	Ein Elternteil ist im Ausland geboren	Erste Generation	Zugewanderte
Mathematik (PISA 2003)	527	508	432	454
Deutsch	517	493	443	437
Englisch	508	507	485	471
Schülerleistungen nach „Umgangssprache in der Familie ist Deutsch“				
	Umgangssprache ist Deutsch		Umgangssprache ist nicht Deutsch	
Mathematik (PISA 2003)	523		434	
Deutsch	523		442	
Englisch	516		482	

Der Migrationshintergrund der Familie führt, wie Tabelle 25.3 ausweist, im Hinblick auf Kompetenzen im Deutschen zu einer ebenso starken Leistungsstreuung wie in Mathematik. Migrationsbedingte Unterschiede im Hinblick auf die Englisch-Leistungen fallen demgegenüber deutlich geringer aus. Insbesondere Kinder der ersten Generation und Zugewanderte erreichen hier höhere Punktzahlen. Dieser Befund wird dadurch unterstrichen, dass für Schülerinnen und Schüler aus Familien, in denen die Umgangssprache nicht Deutsch ist, deutlich höhere Englischkompetenzen als Deutschkompetenzen ermittelt wurden.

7 Dies bestätigt auch die Berechnung der so genannten ‚sozialen Gradienten‘: Ein um eine Standardabweichung höherer höchster sozioökonomischer Status der Familie (HISEI) geht mit einer Steigerung der mathematischen Kompetenz um 38 Punkte einher. Die entsprechenden Kennwerte für Deutsch und Englisch betragen 37 bzw. 35 Punkte.

8 In Anlehnung an PISA 2003 wurde ein vierstufiger Indikator mit den Ausprägungen ‚Kein Migrationshintergrund‘, ‚ein Elternteil im Ausland geboren‘, ‚Erste Generation‘ und ‚Zugewanderte‘ gebildet.

25.6 Struktur- und Prozessmerkmale sozialer Herkunft und Schülerleistungen in sprachlichen Fächern – ein Mediationsmodell

Zur Untersuchung möglicher Wechselwirkungen von Struktur- und Prozessmerkmalen sozialer Herkunft bei der Beeinflussung von Schülerleistungen in Deutsch und Englisch wurden multiple Regressionen herangezogen. Deutsch- und Englisch-Kompetenzen wurden durch je drei Modelle vorhergesagt. Das erste Modell (Modell 1) wurde in PISA 2003 (vgl. Ehmke u.a. 2004) zur Vorhersage der Mathematikleistungen von Schülerinnen und Schülern in Deutschland herangezogen. Es umfasst strukturelle Aspekte sozialer Herkunft. Diese sind der Erwerbstätigkeitsstatus des Vaters, die Familienstruktur, der höchste sozioökonomische Status in der Familie (HISEI), der höchste elterliche Bildungsabschluss (HISCED), kulturelle Besitztümer in der Familie, der Migrationshintergrund sowie der Indikator zur Umgangssprache in der Familie. Modell 2 enthält als sprachbezogene Prozessmerkmale sozialer Herkunft die Mittelwerte der sieben Skalen zur Sprachpraxis im Elternhaus. In Modell 3 schließlich wurden zur Vorhersage von Deutsch- bzw. Englischleistung sowohl strukturelle (Modell 1) wie prozessuale Aspekte (Modell 2) sozialer Herkunft herangezogen. In allen Modellen wurden zudem die numerische kognitive Grundfähigkeit von Schülern sowie Leistungsunterschiede zwischen verschiedenen Bildungsgängen statistisch konstant gehalten. Entsprechende Ergebnisse zeigt Tabelle 25.4.

Strukturmerkmale sozialer Herkunft und Schülerleistung – Der Vergleich zu PISA 2003 (Ergebnisse von Modell 1)

Modell 1 vermag in beiden Fächern substanzielle Anteile von Schülerleistung zu erklären. Statistisch bedeutsame Effekte werden für kulturelle Besitztümer in der Familie, den höchsten sozioökonomischen Status (HISEI) sowie für höchste elterliche Bildungsabschlüsse (HISCED) ermittelt. Hohe Ausprägungen dieser strukturellen Merkmale sozialer Herkunft gehen auch mit höheren Deutsch- und Englisch-Kompetenzen der Schüler einher. Auf Leistungen im Fach Deutsch hat zudem der Tatbestand, dass Deutsch Umgangssprache in der Familie ist, einen positiven Einfluss. Schülerinnen und Schüler mit Migrationshintergrund zeigen hingegen eher geringe Deutschleistungen. Auf Kompetenzen im Fach Englisch hat eine Teilzeitbeschäftigung bzw. Arbeitslosigkeit des Hauptverdieners der Familie einen leicht negativen Effekt.

Tabelle 25.4: Beeinflussung von Schülerleistungen in Deutsch und Englisch durch Struktur- und Prozessmerkmale sozialer Herkunft^A.

	Deutschleistungen			Englischleistungen		
	M 1	M 2	M 3	M 1	M 2	M 3
Strukturmerkmale sozialer Herkunft:						
Erwerbstätigkeitsstatus des Vaters	ns	-	ns	-.03*	-	ns
Familienstruktur (Mutter/Vater alleinerziehend)	ns	-	ns	ns	-	ns
Höchster sozioökonom. Status in der Familie	.07**	-	.07**	.07**	-	ns
Höchster elterlicher Bildungsabschluss	.04*	-	ns	.04**	-	ns
Kulturelle Besitztümer in der Familie	.09**	-	.10**	.08**	-	.04*
Migrationshintergrund der Familie	-.10**	-	-.11**	ns	-	ns
Umgangssprache in der Familie ist Deutsch	.05**	-	.04*	ns	-	ns
Prozessmerkmale sozialer Herkunft:						
Elterneinschätzungen zu ...						
... sprachbezogenen Kenntnissen der Familie	-	.05**	ns	-	.18**	.16**
... Bedeutung der Sprache im Beruf der Mutter	-	ns	ns	-	ns	-.03*
... Bedeutung der Sprache im Beruf des Vaters	-	ns	ns	-	ns	ns
... sprachbez. Unterstützung in der Familie	-	ns	-.11**	-	-.12**	-.11**
... Umgang mit Deutsch / Englisch in der Familie	-	ns	ns	-	.07**	.06**
... Wertschätzung der Sprache in der Familie	-	ns	ns	-	.11**	.10**
... Elterninteresse am Deutsch-/Englischunterricht	-	ns	ns	-	-.04**	-.03*

Anmerkungen: ^AWiedergegeben sind standardisierte Regressionskoeffizienten; ns = nicht signifikant; Signifikanzniveaus: * $p < .05$. ** $p < .01$.

Die Befunde aus Modell 1 vervollständigen das auf Mathematik bezogene Bild aus PISA 2003 wie folgt: Bei der Vorhersage von Mathematikleistungen wurde dort der stärkste Effekt für die sozioökonomische Stellung der Familie (HISEI) ermittelt. Für den Besitz an Kulturgütern und höchste elterliche Bildungsabschlüsse wurden weniger hohe, in etwa gleichwertige Koeffizienten ermittelt. Wie hier wiesen Kinder aus Familien mit hohem sozioökonomischem Status, hohen elterlichen Bildungsabschlüssen und vielen kulturellen Besitztümern in der Familie auch höhere Kompetenzen auf. Mathematikleistungen standen ferner mit der Nicht-Vollzeitbeschäftigung des Hauptverdieners und – abweichend von hier berichteten Befunden – der Familienstruktur in negativem Zusammenhang.

Bei einer Gesamtschau entsprechender Befunde für die Fächer Deutsch, Englisch und Mathematik kann also festgehalten werden: Sprachliche Leistungen werden weniger stark von der sozioökonomischen Stellung der Familie und in höherem Ausmaß von Indikatoren objektivierte kulturellen Kapitals der Familie bestimmt. Der Migrationshintergrund einer Familie hat zwar negative Auswirkungen auf Kompetenzen in Mathematik und Deutsch, steht aber in keinem Zusammenhang zu den Englischkompetenzen (vgl. Kapitel 20).

Prozessmerkmale sozialer Herkunft und sprachliche Leistungen – Ergebnisse von Modell 2

Bei der Vorhersage von sprachlichen Leistungen durch Prozessmerkmale sozialer Herkunft (Modell 2) ergeben sich in beiden Fächern verschiedene Ergebnismuster. Für Leistungen im Deutschen sind die in den Skalen erfassten Einstellungen und Werthaltungen der Eltern fast nicht von Bedeutung. Lediglich das berichtete Ausmaß innerfamiliärer Kenntnisse (Beispielitem: „Ich glaube, in Deutsch wissen wir so gut Bescheid wie die Lehrer unseres Kindes.“) hat einen positiven Einfluss auf Deutsch-Kompetenzen. Anders verhält es sich im Fach Englisch: Positive Elternangaben im Hinblick auf innerfamiliäre Kenntnisse, konstruktiven Umgang und Wertschätzung gehen auf Schülerseite mit höheren Englisch-Kompetenzen einher. Ein Interesse von Eltern am Englischunterricht sowie die grundsätzliche Bereitschaft zu innerfamiliärer Unterstützung (z.B. „Wenn unser Kind in Englisch nicht mehr weiter weiß, helfen wir ihm.“) wird dagegen vornehmlich von Eltern berichtet, deren Kinder eher schlechte Englischleistungen aufweisen. Dieses letzte Ergebnis soll angesichts der oben berichteten Befunde so interpretiert werden, dass tatsächliche sprachbezogene Unterstützung tendenziell eher in Familien mit widrigen Bedingungen sozialer Herkunft und, salopp gesprochen, erst dann erfolgt, „wenn das Kind bereits in den Brunnen gefallen ist“.

Wechselwirkung von Struktur- und Prozessmerkmalen sozialer Herkunft bei der Vorhersage von sprachlichen Leistungen – Ergebnisse von Modell 3

Werden zur Vorhersage von Schülerleistungen Struktur- und Prozessmerkmale sozialer Herkunft simultan herangezogen, ergibt sich im Fach Deutsch ein nahezu unverändertes Bild. Strukturelle Aspekte sozialer Herkunft behalten – mit Ausnahme des höchsten elterlichen Bildungsabschlusses (HISCED) – ihre Effekte bei. Die Sprachpraxis im Elternhaus ist auch bei Kontrolle der sozialen Herkunft nicht prädiktiv für Deutschleistungen. Im Fach Englisch hingegen zeigt sich: Bis auf den Besitz an Kulturgütern – objektiviertes kulturelles Kapital – sind bei Kontrolle von Prozessmerkmalen sozialer Herkunft deren strukturelle Aspekte unerheblich für Schülerleistung. Für Skalen zur Sprachpraxis im Elternhaus werden nach wie vor hochsignifikante, mit Modell 2 konsistente Effekte auf Englischleistung ermittelt.

Die neuentwickelten Indikatoren zur Erfassung sprachbezogener Elterneinstellungen und Werthaltungen von Eltern im Sinne von Prozessmerkmalen sozialer Herkunft liefern für das Fach Deutsch also keinen über den der in anderen Studien bewährten Strukturmerkmale sozialer Herkunft hinausgehenden Informationsgewinn. Die Effekte von Strukturmerkmalen sozialer Herkunft auf Englischleistung jedoch werden fast vollständig durch die Skalen zur Sprachpraxis im Elternhaus, also durch Prozessmerkmale sozialer Herkunft, mediert (vgl. Baron/Kenny 1986).

Soziale Herkunft, Sprachpraxis und Lernzuwächse in Deutsch und Englisch

Das längsschnittliche Design von DESI erlaubt die Beantwortung der Fragestellung, inwieweit strukturelle und prozessuale Aspekte sozialer Herkunft prädiktiv für die Veränderung von Kompetenzen in Deutsch und Englisch sind. Dazu wurden die Analysen aus Tabelle 25.4 für die erfassten Kompetenzzuwächse repliziert.

Im Fach Deutsch (betrachtet wurde der Indikator zum Leseverstehen) zeigt sich bei vergleichsweise geringen Zuwächsen, dass lediglich der kulturelle Besitz in der Familie vorhersagekräftig für Leistungsveränderungen ist. Je mehr Kulturgüter eine Familie besitzt, umso höher ist dieser Zuwachs. Für Englischleistungen ergibt sich auf der Grundlage des eingesetzten C-Tests folgendes Bild: Lernzuwächse sind unabhängig von allen Strukturmerkmalen sozialer Herkunft. Die von Eltern berichtete Wertschätzung der Sprache hat dagegen einen signifikant positiven Einfluss auf Leistungszuwächse in diesem Fach. Diese skizzenartigen Befunde zur Determination von Lernzuwächsen innerhalb des Untersuchungszeitraumes von DESI unterstreichen ein weiteres Mal die differentielle Wirksamkeit von Struktur- und Prozessmerkmalen sozialer Herkunft bei der Beeinflussung von Deutsch- und Englischleistungen.

25.7 Zusammenfassung

Mit den Analysen dieses Kapitels sollte zunächst im Hinblick auf Strukturmerkmale sozialer Herkunft ein möglichst hohes Maß an Vergleichbarkeit zu entsprechenden Befunden von PISA 2003 hergestellt werden. Weiterhin war beabsichtigt, erstmals sprachbezogene Prozessmerkmale sozialer Herkunft zu beschreiben und ihr Verhältnis zu strukturellen Bedingungen sozialer Herkunft zu untersuchen. Am Ende stand die Fragestellung, inwieweit der Einfluss von Strukturmerkmalen sozialer Herkunft auf Schülerleistungen in Deutsch und Englisch durch die Ausprägungen von Prozessmerkmalen sozialer Herkunft transportiert wird.

Der Vergleich mit den in PISA 2003 gewonnenen Ergebnissen lässt sich wie folgt zusammenfassen: Im Hinblick auf die Erwerbstätigkeit der Eltern, die Familienstruktur, den höchsten sozioökonomischen Status, höchste elterliche Bildungsabschlüsse und kulturellen Besitz in der Familie als *Strukturmerkmale sozialer Herkunft* wurden in der Elternschaft von DESI mit PISA 2003 in hohem Maße vergleichbare Bedingungsgefüge ermittelt. Auf deskriptiver Ebene ergeben sich zwischen strukturellen Bedingungen sozialer Herkunft und mittleren Schülerleistungen in Deutsch, Englisch und Mathematik durchaus differentielle Zusammenhänge. Für Kinder aus Familien, die über unterschiedlich viel objektiviertes kulturelles Kapital (kulturellen Besitztümern) verfügen, zeigen sich in sprachlichen Fächern höhere Leistungsstreuungen als in Mathematik. Unterschiedliche Bedingungen im Hinblick auf die sozioökonomische Stellung der Familie gehen dagegen mit weniger hohen Unterschieden in Schülerleistungen in Deutsch und Englisch einher als in PISA 2003 (Mathematik). Der Migrationshintergrund von Familien und der Tatbestand,

dass zu Hause nicht Deutsch als Umgangssprache gesprochen wird, hat auf fremdsprachliche Leistungen geringere Auswirkungen als auf Deutsch- und Englisch-Kompetenzen. Diese Befunde bleiben auch dann erhalten, wenn Strukturmerkmale sozialer Herkunft regressionsanalytisch zu Schülerleistungen in sprachlichen Fächern in Beziehung gesetzt werden und dabei für Unterschiede im Hinblick auf kognitive Grundfähigkeiten von Schülern und verschiedene Leistungsniveaus in den hier betrachteten Bildungsgängen kontrolliert wird.

Zur Beschreibung *sprachbezogener Prozessmerkmale sozialer Herkunft* wurden auf die Fächer Deutsch und Englisch bezogene Werthaltungen und Einstellungen von Eltern (Sprachpraxis im Elternhaus) sowie von Eltern berichtete Häufigkeiten tatsächlichen sprachbezogenen Unterstützungsverhaltens herangezogen. Im Hinblick auf das Fach Deutsch wurde dabei eine förderlichere Sprachpraxis im Elternhaus als für das Fach Englisch ermittelt. Diese äußert sich in höheren selbstberichteten Kenntnissen, einer größeren Bedeutung von Deutsch im Beruf von Mutter und Vater, einem konstruktiveren Umgang mit dieser Sprache und einer höheren prinzipiellen Bereitschaft, das Kind bei Fehlern und Verständnislücken zu unterstützen. Die auf Deutsch und Englisch bezogenen Werthaltungen und Einstellungen von Eltern kovariieren erwartungsgemäß mit Strukturmerkmalen sozialer Herkunft, so dass hohe Ausprägungen der Letzteren auch mit positiverer Sprachpraxis im Elternhaus einhergehen. Dies aber äußert sich *nicht* in der tatsächlichen sprachbezogenen Unterstützung durch die Eltern. Derartige Unterstützung, etwa in Form von Hausaufgabenhilfe, erfolgt im Gegenteil eher in Familien aus widrigen strukturellen Bedingungen sozialer Herkunft und wenn Eltern unzufrieden mit den Leistungen ihres Kindes sind. Dieser offensichtlich kompensatorische Charakter tatsächlicher sprachbezogener Unterstützung ist unabhängig von Strukturmerkmalen sozialer Herkunft.

Im Hinblick auf die *Wechselwirkung von Struktur- und Prozessmerkmalen sozialer Herkunft bei der Beeinflussung von Schülerleistungen in Deutsch und Englisch* zeigte sich: Leistungen im Deutschen sind fast ausschließlich durch strukturelle Aspekte sozialer Herkunft, also zeitstabile Bedingungen wie den sozioökonomischen Status oder das elterliche Bildungsniveau, beeinflusst. Prozessmerkmale sozialer Herkunft im Sinne sprachbezogener Einstellungen und Werthaltungen von Eltern (Sprachpraxis im Elternhaus) sind hier nahezu bedeutungslos. In Bereich Englisch ergibt sich ein umgekehrtes Befundmuster: Die Effekte des sozioökonomischen Status, der elterlichen Bildungsabschlüsse und des objektivierten kulturellen Kapitals (Strukturmerkmale) verschwinden, sobald für die Sprachpraxis im Elternhaus (Prozessmerkmale) kontrolliert wird. Die Beeinflussung der Englischleistungen durch Strukturmerkmale wird hier offensichtlich vollständig über die Ausprägungen der Prozessmerkmale transportiert. Bei Schülern aus Familien, deren Eltern hohe fachbezogene Kenntnisse, eine hohe Wertschätzung des Faches Englisch und einen positiven Umgang mit dieser Sprache berichten, werden dabei auch höhere Kompetenzen ermittelt.

Literatur

- Baron, R. M./Kenny, D. A. (1986): The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. In: *Journal of Personality and Social Psychology* 51, S. 1173-1182.
- Baumert, J./Schümer, G. (2001): Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K. J./Weiß, M. (Hrsg.): *PISA 2000 - Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich, S. 323-397.
- Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K. J./Weiß, M. (Hrsg.) (2001): *PISA 2000 - Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich.
- Baumert, J./Watermann, R./Schümer, G. (2003): Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs: Ein Institutionelles und individuelles Mediationsmodell. In: *Zeitschrift für Erziehungswissenschaft* 6, H. 1, S. 46-72.
- Bernstein, B. (1977): *Beiträge zu einer Theorie des pädagogischen Prozesses*. Frankfurt am Main: Suhrkamp.
- Bourdieu, P. (1983): Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In: Kreckel, R. (Hrsg.): *Soziale Ungleichheiten. Soziale Welt, Sonderband 2*. Göttingen: Schwartz, S. 183-198.
- Coleman, J.S. (1966): *Equality of Educational Opportunity*. Washington, D.C.: U.S. Department of Health, Education, and Welfare.
- Coleman, J.S. (1988): Social Capital in the Creation of Human Capital. In: *American Journal of Sociology* 94, Supplement, S. 95-120.
- Ehmke, T./Hohensee, F./Heidemeier, H./Prenzel, M. (2004): Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In: Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M./Pekrun, R./Rolff, H.-G./Rost, J./Schiefele, U. (Hrsg.): *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann, S. 226-253.
- Ehmke, T./Siegle, T./Hohensee, F. (2005): Soziale Herkunft im Ländervergleich. In: Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M./Pekrun, R./Rost, J./Schiefele, U. (Hrsg.): *PISA 2003. Der zweite Vergleich der Länder in Deutschland – Was wissen und können Jugendliche*. Münster: Waxmann, S.233-264.
- Ganzeboom, H. B. G./de Graaf, P. M./Treiman, D. J./de Leeuw, J. (1992): A standard international socio-economic index of occupational status. In: *Social Science Research* 21, S. 1-56.
- Helmke, A./Schrader, F. W./Hosenfeld, I. (2004): Elterliche Lernunterstützung und Schulleistungen Ihrer Kinder. In: *Bildung und Erziehung* 57, S. 251-277.
- Köller, O./Watermann, R./Trautwein, U./Lütke, O. (Hrsg.) (2004): *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*. Opladen: Leske & Budrich.
- Organisation for the Economic Co-Operation and Development (OECD) (1999): *Classifying educational programmes. Manual for ISCED-97 implementation in OECD countries*. Paris: OECD.
- Portes, A. (1998): Social Capital: Its Origins and Applications in Modern Sociology. In: *Annual Review of Sociology* 24, S. 1-24.
- Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M./Pekrun, R./Rolff, H.-G./Rost, J./Schiefele, U. (Hrsg.) (2004): *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Rolff, H.-G. (1967): *Sozialisation und Auslese durch die Schule*. Heidelberg: Quelle und Meyer.
- Westat (2002): *WesVar 4.2 User's Guide*. Westat: Rockville, MD.

Andreas Helmke / Eckhard Klieme

26 Unterricht und Entwicklung sprachlicher Kompetenzen

DESI bietet infolge seiner Anlage als Zweipunktmessung sprachlicher Kompetenzen und motivationaler Orientierungen die Chance, nach Bedingungen von Veränderungen zu fragen und geht damit über die Erkenntnismöglichkeiten querschnittlich angelegter Studien wie PISA 2000, IGLU oder MARKUS deutlich hinaus. Eine zentrale Fragestellung von DESI ist, welche Merkmale von Lehrpersonen und des Fachunterrichts für den Kompetenzzuwachs eine Rolle spielen. Wir skizzieren zunächst einige grundlegende theoretische und methodische Fragen der Unterrichtsforschung. Anschließend geben wir einen Überblick über die wichtigsten Analysestrategien, mit denen bei DESI die Wirksamkeit des Deutsch- und Englischunterrichts untersucht wird.

26.1 Ein Rahmenmodell der Wirksamkeit des Unterrichts

Grundlage der Erhebungen bei DESI ist, wie in Kapitel 1 erläutert, ein allgemeines Rahmenmodell, das einerseits dem Ziel dient, die Komplexität der Bedingungen und der Wirksamkeit des Unterrichts zu veranschaulichen; andererseits soll klargestellt werden, zu welchen der Bedingungs-, Erklärungs- und Wirkungsblöcke seitens DESI Informationen erhoben wurden, welche Ausschnitte aus den zahlreichen unterrichtsbezogenen Fragestellungen Gegenstandsbereich des vorliegenden Buches sind und welche für spätere vertiefende Auswertungen vorgesehen sind. Die von Helmut Fend (1981) übernommene Grundidee des Angebots-Nutzungs-Modells ist die, dass der Unterricht als ein Angebot zu verstehen ist, das seitens der Adressaten – der Schülerinnen und Schüler – genutzt werden kann. Ob, wie effizient und wie nachhaltig dies erfolgt, hängt vom Lernpotenzial der Schüler, vom Kontext und von verschiedenen Unterrichtsmerkmalen ab. Im Zentrum der Kapitel 26 bis 33 stehen die Faktoren, die in Abbildung 26.1 oben links aufgeführt sind: Merkmale der Lehrperson und des Unterrichts.

im laufenden Schuljahr, Unterrichtsausfall, Unterrichtsvertretungen, Zeitverlust durch Raumwechsel etc.) und der Schülerinnen und Schüler (individuelle Fehlzeiten infolge von Krankheit u.a.) vor, die im vorliegenden Buch jedoch aus Platzgründen nicht thematisiert werden. Das Gleiche gilt für die Qualität des *Lehr-Lern-Materials*; hierzu liegen seitens der Lehrkräfte Einschätzungen der Qualität des verwendeten Lehrwerks vor.

Als Zielkriterien (*Wirkungen*) fungieren bei den folgenden Auswertungen primär Aspekte der Sprachkompetenz im Deutschen und Englischen. Aussagen über Veränderungen im Laufe der neunten Jahrgangsstufe können jedoch auch zu anderen Schülermerkmalen getroffen werden, wenn sie sowohl zu Beginn als auch am Ende der neunten Jahrgangsstufe erhoben wurden. Dazu gehören die Lernmotivation, soziales und akademisches Selbstkonzept wie auch allgemeine Lernstrategien.

Analysen zur Wirksamkeit des Unterrichts, wie sie vor allem in den Kapiteln 28 und 32 vorgestellt werden, haben auch Faktoren in Rechnung zu stellen, die dem Unterricht vorgelagert sind und/oder ihn begleiten, d.h. den kulturellen, familiären und persönlichen Hintergrund und die eigenen Lernaktivitäten der Jugendlichen (zu Faktoren der Schulebene, die hier nicht abgebildet ist, vgl. Kapitel 34). Solche Faktoren wurden in den Kapiteln 19 bis 25 bereits ausführlich diskutiert und sollen hier nur kurz im Blick auf ihre Relevanz für Unterrichtsprozesse erwähnt werden.

Familie. Strukturelle und Prozess-Merkmale der Familie gehen – auf Individual-ebene und aggregiert auf Schulklassenebene – teilweise in den Komplex „Sozialer und kognitiver Hintergrund“ mit ein. Vorstellungen zum komplexen Zusammenwirken von familiären, individuellen und unterrichtlichen Merkmalen bilden den Fokus des Kapitels 25.

Individuelles Lernpotenzial. Dies umfasst die wesentlichen biologischen und psychologischen Schülermerkmale, die ihrerseits durch Genetik sowie familiäre und schulische Sozialisation bedingt sind. Neben den Sprachkompetenzen selbst zielt DESI insbesondere auf solche kognitiven, volitionalen und motivationalen Schülermerkmale, die sowohl Bedingungen des Lernens und Leistens sind als auch Zielkriterien des Unterrichts: Leseinteresse, Lerninteresse und Lernstrategien.

Lernaktivitäten. Unterricht kann sich nur insofern auf den Leistungszuwachs bzw. die Kompetenzentwicklung auswirken, als individuelle Lernaktivitäten angeregt, gestützt und aufrechterhalten werden. Lernaktivitäten finden während des Unterrichts statt, aber auch die Hausaufgaben und außerschulische Erfahrungsfelder bieten Gelegenheiten für fachrelevantes Lesen und Lernen.

Kontext. Der Kontext – d.h. die zu Beginn der neunten Jahrgangsstufe vorgefundenen Eingangs- und Rahmenbedingungen, insbesondere die Klassenzusammensetzung und das Vorkenntnisniveau – spielten für die Analysen eine überragende Rolle, auf die in Abschnitt 26.4 unter dem Stichwort „Hintergrundmodell“ noch eingegangen wird. Hier nur soviel:

Die *historischen und kulturellen Rahmenbedingungen* sind diejenigen, die DESI im Deutschland im Jahr 2003 vorgefunden hat – interessante Querbeziehungen können sich in der Zukunft durch gelegentliche *follow up*-Messungen in Deutschland

ergeben. Der *regionale Kontext* spielt bei DESI keine eigenständige Rolle, obgleich Daten zu Stadt/Land, Einzugsgebiet etc. erhoben wurden. Umso entscheidender sind die *Bildungsgänge* und – auf der Schulebene – *Schularten*. An einigen Stellen gehen wir der Frage nach, ob der Einfluss von Merkmalen der Unterrichtsqualität – z.B. das Ausmaß, in dem im Unterricht Englisch gesprochen wird – vom *didaktischen Kontext* moderiert wird, etwa in der Weise, dass die Folgen der Verwendung des Deutschen als Unterrichtsprache davon abhängt, ob dies im Rahmen des Unterrichtsgesprächs oder in anderen Phasen erfolgt. Schließlich wurden bei DESI aus Schülersicht Angaben zum *klimatischen Kontext* in der Klasse („Klassenklima“, „Klassengeist“, „Schüler-Schüler-Verhältnis“) erhoben.

Am Beispiel des Klassenklimas, d.h. der von den Beteiligten erlebten Qualität sozialer Beziehungen in der Klasse, lassen sich auch die Grenzen eines schematisierten Wirkungsdiagramms zeigen, wie es Abbildung 1 darstellt. Die sozialen Beziehungen (Lehrer – Schüler, Schüler – Schüler) bestehen ja nicht außerhalb von Unterrichtsprozessen und Lernaktivitäten, sondern sie werden von diesen geformt, lassen sich auch als Merkmale der unterrichtlichen Interaktion verstehen. Einem modernen, „konstruktivistischen“ Verständnis pädagogischer Prozesse folgend stellt die Lernumgebung in Schule und Unterricht ein „Angebot“ dar, das von den Lernenden aktiv „genutzt“ werden muss (Fend 2006; Helmke 2004). Das Bild einer linearen „Wirkungskette“ von Kontext- über Prozess- zu Ertragsmerkmalen würde diesem Verständnis nicht gerecht. DESI bildet vielmehr die Wechselwirkung von Unterricht und Lehrerhandeln einerseits, Schülermerkmalen andererseits über die neunte Jahrgangsstufe hinweg ab.

26.2 Prozessqualität des Unterrichts

Unterrichtsforschung und Lehr-Lern-Forschung haben in gut 100 Jahren empirischer Forschung einen beachtlichen Korpus an unterrichtsrelevantem Wissen zusammengetragen (für eine Übersicht vgl. Helmke 2004; Pauli/Reusser 2006; Shuell 1996). Dazu kommen Forschungen zur Lehrerexpertise und Lehrerprofessionalität (Berliner 1995; Bromme 1997; Lipowsky 2006), die sich allerdings mehr für Prozesse der Planung, des Denkens, der Wissensorganisation von Lehrern und weniger für „outcomes“ interessiert haben.

Der Ertrag der Unterrichtsforschung ist vielfach in Form von allgemeinen (oder „generischen“), d.h. nicht auf ein bestimmtes Fach bezogenen Prinzipien oder Gütekriterien des Unterrichts zusammengefasst worden. Weitgehenden Konsens gibt es inzwischen darüber, dass es den idealen Unterrichtsstil, die optimale Unterrichtsmethode nicht gibt und angesichts der Vielfalt an Bildungszielen, der Unterschiedlichkeit der Lernvoraussetzungen und der fachlichen Inhalte und Settings auch nicht geben kann. Die alten Debatten über die angebliche Über- oder Unterlegenheit bestimmter Unterrichtsstile (wie Offener Unterricht vs. Direkte Instruktion) sind inzwischen obsolet. Das Gleiche gilt für die Gegenüberstellung frontaler lehrerzentrierter Methoden einerseits und konstruktivistisch orientierter und

gruppenbasierter Methoden andererseits (Gudjons 2003). Was es dagegen gibt, sind zusammenfassende Erkenntnisse über allgemeine, fachübergreifende Gütekriterien des Fachunterrichts (Helmke 2004):

- Effiziente Klassenführung und Zeitnutzung
- Lernförderliches Unterrichtsklima
- Vielfältige Motivierung
- Strukturiertheit und Klarheit
- Wirkungs- und Kompetenzorientierung
- Schülerorientierung, Unterstützung
- Förderung aktiven, selbstständigen Lernens
- Angemessene Variation von Methoden und Sozialformen
- Konsolidierung, Sicherung, Intelligentes Üben
- Passung von Unterrichtsangebot und Lernvoraussetzungen

Diese und ähnliche Übersichten (Brophy 2000; Meyer 2004) sind eine Orientierungshilfe und stellen zugleich das Gerüst von „Rating“-Bögen zur Unterrichtsbeobachtung und -bewertung dar, wie sie in inhaltlich sehr ähnlicher Form im Rahmen der externen Schulevaluation eingesetzt werden (vgl. Helmke 2007; Helmke/Schrader in Druck). Für die wissenschaftliche Erforschung des Unterrichts sind diese Listen jedoch von eingeschränktem Wert. Zur Vermeidung von Fehlinterpretationen und Kurzschlüssen ist auf folgende Einschränkungen hinzuweisen:

Erstens umfasst die empirische Basis im Wesentlichen Studien aus dem *angloamerikanischen* Bereich; in Deutschland hat sich die empirische Unterrichtsforschung gerade erst entfaltet (Klieme 2006; Wellenreuther 2005). Gerade DESI kann zu einer Verbreiterung und Vertiefung empirisch fundierten Wissens über die Unterrichtswirksamkeit beitragen.

Zweitens beziehen sich die genannten Gütekriterien auf *Unterricht im Allgemeinen*, sehen also von der speziellen Domäne „Deutschunterricht“ bzw. „Fremdsprachenunterricht“ ab. Dieses Defizit ist bereits vor einem Vierteljahrhundert von Achtenhagen zu Recht kritisiert worden, steht aber noch heute auf der Agenda der Unterrichtsforschung (Baumert u.a. 2004). Wo Unterrichtsforschung fachdidaktisch fundiert ist, hat sie sich bislang überwiegend mit Mathematik und Naturwissenschaften befasst. Aus diesem Grund wurden bei DESI neben fachübergreifenden („generischen“) auch für das Unterrichtsfach Deutsch bzw. Englisch spezifische Instrumente entwickelt.

Drittens finden sich in empirisch vorfindbaren unterrichtsbezogenen Einschätzungen – z.B. in fragebogenbasierten Erhebungen der Unterrichtswahrnehmung durch Schüler – in aller Regel andere, weniger differenzierte *Urteilsstrukturen* als in den o.g. Konstruktlisten. Das Problem wird noch dadurch verschärft, dass Erhebungen des Unterrichts aus unterschiedlichen Perspektiven (z.B. Schüler – Lehrpersonen – Dritte) zu ganz unterschiedlichen Dimensionen „der“ Unterrichtsqualität führen können und jeweils spezifische Stärken und Schwächen aufweisen, wie Clausen (2002) nachgewiesen hat. Ebenso erhält man unterschiedliche Urteilsstrukturen in Abhängigkeit von der zeitlichen Erstreckung des zu beurteilenden Unterrichts (ku-

mulativer Ansatz: Beurteilung überdauernder Merkmale „des“ Unterrichts einer Lehrperson, versus punktuell: Beurteilung einer konkreten Unterrichtsstunde.)

Eine Stoßrichtung der modernen Unterrichtsforschung ist deshalb die Identifikation von übergeordneten Faktoren (Ditton 2006), insbesondere von empirisch unterscheidbaren, voneinander hinreichend unabhängigen und sparsamen Kernprinzipien der Qualität des Fachunterrichts¹. Hierzu wurde basierend auf der TIMS-Studie und TIMSS-Videostudie für den Mathematikunterricht ein dreidimensionales Modell entwickelt (Klieme 2002), das die folgenden Dimensionen umfasst: schülerorientiertes Unterrichtsklima, kognitive Aktivierung und effiziente Klassenführung. Gestützt auf kognitions-, motivations- und instruktionspsychologische Annahmen wird erwartet, dass die erstgenannte Grunddimension primär für Motivationsförderung im Unterricht relevant ist, die zweite Grunddimension für fachliches Verständnis und Wissenserwerb, während die dritte eine notwendige Voraussetzung für kognitive wie motivationale Prozesse bildet. Dieses Modell bewährte sich bei der Erklärung der Wirksamkeit von Mathematikunterricht wie auch in anderen Studien (Klieme/Rakoczy 2003; Klieme/Rakoczy/Lipowsky/Ratzka 2006; Brunner u.a. 2006), kann allerdings – aus Gründen der Fachspezifität – nicht ohne weiteres auf den Deutsch- und Englischunterricht übertragen werden. Dies gilt insbesondere für die *kognitive Aktivierung*, die in Mathematik als Strukturiertheit und Tiefe der logischen Argumentation unerlässlich ist, angesichts anderer (insbesondere kommunikativer) Lernziele im Deutsch- und Englischunterricht jedoch neu interpretiert werden muss (vgl. Klieme 2006). Wir werden deshalb die Frage der Dimensionalität der Unterrichtsqualität aus Schülersicht separat für die beiden Fächer empirisch überprüfen.

26.3 Methodische und theoretische Probleme der Forschung zur Unterrichtswirksamkeit

Fundierte Analysen der Unterrichtswirksamkeit müssen eine Reihe von Problemen adäquat berücksichtigen. Die wichtigsten Punkte sind aus unserer Sicht die folgenden:

Universelle und inkrementelle Wirksamkeit von Schule und Unterricht. Was den Einfluss der Schule und des Unterrichts auf die Leistungsentwicklung anbelangt, so trifft man vielfach auf ein fundamentales Missverständnis: Aus der Stabilität der *Rangordnung* schulklassenbezogener oder individueller Leistungsprofile über die Zeit, also aus der geringen Veränderbarkeit individueller Unterschiede bei kognitiven und persönlichen Merkmalen, wie es sich auch bei DESI für die sprachlichen Kompetenzen findet, wird häufig geschlossen, Schule und Unterricht seien faktisch wirkungslos. In der Tat zeigt die Forschung, dass viele Merkmale schulischer Leistungen bereits ab der Grundschule eine enorme Stabilität aufweisen, wie sich beispielsweise in der Längsschnittstudie SCHOLASTIK (Weinert/Helmke

1 Um eine Vielfalt von Kategorien auf eine übersichtliche Anzahl von Dimensionen zu reduzieren, eignen sich Methoden der zweiebenenanalytischen Faktorenanalyse mit Hilfe des Programms Mplus (Muthén/Muthén 2003).

1997) herausgestellt hat. Richtet man dagegen den Blick auf *Lernfortschritte*, dann kommt der Schule eine überragende Bedeutung zu. Dies führte Weinert (1992) zu der These: „Was Schulen für die geistige Entwicklung der Kinder leisten, zeigt sich dann am deutlichsten, wenn es keine gibt“ (S. 8/9). Kontrollierte Vergleichsstudien zur Entwicklung von Kindern mit und ohne Schulbesuch in Ländern wie Peru, Marokko und Guatemala machen deutlich, dass die Wirkung des Schulunterrichts nicht auf den Erwerb grundlegender Kulturtechniken wie Lesen, Schreiben und Rechnen oder die systematische Wissensvermittlung beschränkt bleibt. „Die Schule ist nämlich darüber hinaus von grundlegender Bedeutung für die Entwicklung der höheren Formen des Lernens und des Denkens. Das Verständnis für komplexe Zusammenhänge, der Umgang mit abstrakten Symbolen und die Verwendung reflexiver Strategien des Problemlösens haben stets ein Minimum an schulischer Unterweisung zur notwendigen Voraussetzung“ (Weinert 1992, S. 9).

Mehrebenencharakter der Datenbasis. Die Analysen zur Unterrichtswirksamkeit müssen dem Sachverhalt Rechnung tragen, dass die unterrichtlichen Angaben zur Qualität des Unterrichts ein- und derselben Lehrkraft innerhalb der Klasse streuen, dass also die Schülerwahrnehmungen des Unterrichts in einer Klasse mehr oder weniger *geteilt* sind. Die hierarchische Datenstruktur (Schüler- und Klassenebene) zu ignorieren und der Einfachheit halber nur mit Klassenmittelwerten zu operieren, wäre ein folgenreicher methodischer Fehler. Deshalb werden bei DESI alle Analysen zum Thema „Unterricht“ mehrebenenanalytisch durchgeführt.

Schwache, aber inhaltlich relevante Effekte. Weil Unterricht in Ländern mit einheitlichen Curricula und starker pädagogisch-professioneller Tradition – sogar über Staaten hinweg – recht festen Mustern folgt („Skripts“; vgl. Pauli/Reusser 2006), variieren Unterrichtsmerkmale nur begrenzt. Schon deshalb ist es schwer, in Survey-Studien wie DESI bedeutsame Effekte von Unterrichtsmerkmalen auf die Leistungs- und Motivationsentwicklung der Schülerinnen und Schüler zu finden. Zudem lässt sich die individuelle Nutzung der Lernangebote hier nur begrenzt ermitteln und in den Analysen berücksichtigen. Diese Gründe führen dazu, dass Unterrichtseffekte in empirischen Studien wie DESI nach den gängigen Maßstäben der empirischen Sozialforschung meist schwach ausfallen, jedenfalls schwächer als die Effekte individueller Voraussetzungen wie etwa des Vorwissens. Pädagogisch aber sind die Unterrichtseffekte von zentraler Bedeutung, denn sie allein geben Hinweise zur Verbesserung der Lernumgebung in den verschiedenen Fächern.

26.4 Analysestrategien und -probleme

In den folgenden Kapiteln werden gelegentlich Maße des Zusammenhangs zwischen Unterrichtsmerkmalen und Merkmalen der Lehrperson einerseits und verschiedenen Maßen der sprachlichen Kompetenz andererseits berichtet. Dabei werden unterschiedliche Arten der Korrelation mit unterschiedlicher Bedeutung unterschieden. Dies soll im Folgenden kurz veranschaulicht werden.

a) Die *einfache Korrelation* entspricht der Fragestellung „Ist eine höhere Ausprägung des Unterrichts- oder Personenmerkmals mit einem höheren Niveau der Deutsch- bzw. Englischleistung verbunden?“ Vorab sei festgehalten: Einfache Korrelationen, wie sie bei nicht-experimentellen Surveystudien wie DESI anfallen, gestatten grundsätzlich keine Aussagen über Wirkungszusammenhänge und sind deshalb stets sehr vorsichtig zu interpretieren. Wenn wir sie hier trotzdem gelegentlich berichten, dann lediglich mit dem Ziel, am Kontrast zwischen einfachen und partiellen Korrelationen deutlich zu machen, welche Sackgassen und Artefakte nahe liegen, wenn einfache Korrelationen kausal missinterpretiert werden.

Darüber hinaus messen Korrelationen ausschließlich *lineare* Zusammenhänge zwischen Merkmalen des Unterrichts bzw. der Lehrperson und der Kompetenz(entwicklung) auf Seiten der Schülerinnen und Schüler. Eventuell vorhandene *nicht-lineare* Beziehungen bleiben mit Hilfe korrelativer Analyseverfahren unentdeckt (z.B. eine umgekehrt U-förmige Beziehung, wenn sich die beste Kompetenzentwicklung bei einer *mittleren* Ausprägung eines Unterrichtsqualitätsmerkmals zeigt, während sehr niedrige und sehr hohe Ausprägungen gleichermaßen ungünstig sind). Aus diesem Grunde prüfen wir, wo es theoretisch nahe liegt, mit Hilfe regressionsanalytischer Methoden, ob lineare durch nichtlineare Zusammenhänge überlagert werden.

b) *Partielle Korrelationen*: Hier werden Posttest (im Fall von DESI: die Leistungen am Ende der neunten Jahrgangsstufe) und Unterricht um den Einfluss der Eingangsbedingungen bereinigt². Dies ist in Abbildung 26.2 dargestellt. Der Pfad a repräsentiert die Wirksamkeit eines Unterrichtsmerkmals auf den Posttest, wenn man dabei in Betracht zieht, dass sowohl der Posttest (Pfade b und c) als auch der Unterricht selbst (Pfade d und e) vom anfänglichen Kompetenzniveau sowie von der Klassenzusammensetzung abhängig sein können.

2 Alternativ dazu wird in manchen Analysen auch angegeben, welche Rolle Unterrichtsmerkmale für die *Varianz des Zuwachses* (z.B. der Hörverstehensleistung oder des Lerninteresses) spielen. Die dabei resultierenden Schätzungen der Wirksamkeit des Unterrichts fallen deutlich höher aus (Lanahan et al. 2005)

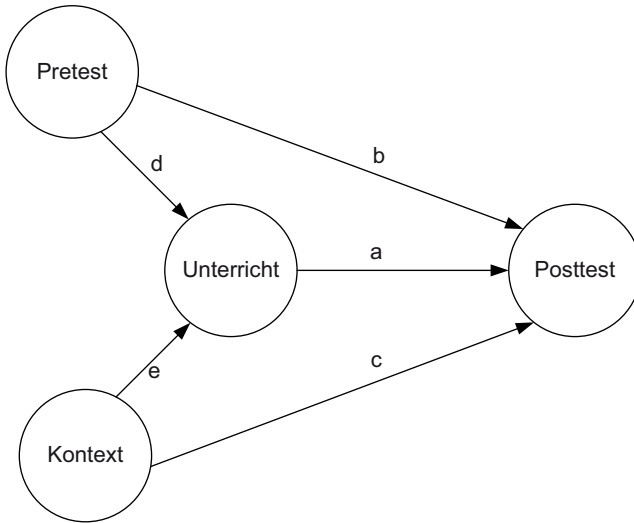


Abbildung 26.2: Unterricht als Resultante (des Kontextes, des Vorkenntnisniveaus) und als Steuergröße für den Posttest.

Die partielle Korrelation trägt der Tatsache Rechnung, dass die Qualität des Unterrichts nicht nur „unabhängige Variable“ ist, sondern selbst von vorgefundenen Rahmenbedingungen, insbesondere vom Vorkenntnisniveau (oder anderen Merkmalen, deren Veränderung gemessen werden soll) abhängen kann. Je nach Klassenzusammensetzung kann ein anspruchsvoller, die Unterrichtszeit fachlich nutzender Unterricht begünstigt oder auch behindert werden.

Allerdings ist dabei zu berücksichtigen, dass Kontext (in Gestalt der Eingangsbedingungen, z.B. des anfänglichen Kompetenzniveaus) und Unterrichtsqualität in einem Verhältnis kontinuierlicher wechselseitiger, also bidirektionaler Beeinflussung stehen, und zwar auf Mikroebene (innerhalb einer Unterrichtsstunde) wie auf Makroebene (im Laufe eines Schuljahres). Das heißt z.B.: Das bei DESI zum „Hintergrund“ gehörende sprachliche Kompetenzniveau oder das Lerninteresse der Schülerinnen und Schüler am Anfang der neunten Jahrgangsstufe ist seinerseits Ergebnis der Qualität des vorangegangenen Unterrichts, der vormaligen Kontextbedingungen usw. Dies gilt nur dann nicht, wenn eine Lehrkraft (a) eine Klasse neu übernimmt und (b) ein bislang noch nie unterrichtetes Fach lehrt. Aus alledem folgt, dass man Zusammenhänge zwischen Kontext und Unterrichtsqualität beachten, aber nur mit Vorbehalt unidirektional interpretieren sollte, auch wenn sie – wie bei DESI – zeitlich nacheinander erhoben wurden. Im Übrigen sollte man bei der Interpretation von Zusammenhängen zwischen Klassenzusammensetzung und Unterrichtsqualität den Ausdruck „Adaptivität“ vermeiden, da dies suggeriert, die je nach günstiger oder ungünstiger Klassenzusammensetzung unterschiedlichen Lehr- und Managementstrategien der Lehrpersonen seien das Ergebnis einer bewusst-rationalen und intentional den Kontext explizit berücksichtigenden Unterrichtsplanung.

Dies kann der Fall sein oder auch nicht – DESI kann dazu keine empirisch fundierten Aussagen machen.

c) Das „*Hintergrundmodell*“. Bei den auf die Wirksamkeit des Unterrichts bezogenen Zweiebenenanalysen von DESI wird auf Klassen- und auf Individualebene standardmäßig ein Set von Kontextvariablen („*Hintergrund*“) aus den Beziehungen zwischen Unterrichtsqualität und Kriterium herausgerechnet, der die folgenden Merkmale umfasst: Bildungsgang (nur Klassenebene), Geschlecht, Erstsprache und kognitive Grundfähigkeit (mit dem KFT erfasst). Diese Strategie dient zwei Zielen: zum einen der Vermeidung von Artefakten (wenn z.B. der Kompetenzzuwachs in Wirklichkeit nicht dem Unterricht, sondern einer günstigen oder ungünstigen Klassenzusammensetzung geschuldet ist), zum anderen der Vergleichbarkeit der Ergebnisse über verschiedene Fächer und Analysen hinweg.

Es soll jedoch an dieser Stelle nicht verschwiegen werden, dass diese Strategie auch ihre Kosten hat: Der Blick auf bildungsgangübergreifende Zusammenhangsmuster unter Auspartialisierung der Bildungsgangszugehörigkeit maskiert möglicherweise ernstzunehmende, bildungssoziologisch und pädagogisch-psychologisch interessante bildungsgang- und schulartspezifische Ergebnismuster (Stichwort „Schulform als differentielles Entwicklungsmilieu“, vgl. Baumert/Trautwein/Artelt 2003). Zudem sind kognitive Grundfähigkeit, Sozialschichtzugehörigkeit und Bildungsgang im deutschen Schulsystem aufgrund der Selektionsprozesse beim Übergang in die Sekundarstufe so stark korreliert, dass ihre gleichzeitige Berücksichtigung im „*Hintergrundmodell*“ zu statistischen Problemen führt (Multikollinearität)³. Dazu kommt das Problem, dass die kognitiven Grundfähigkeiten nicht zum ersten, sondern erst zum zweiten Messzeitpunkt erhoben werden konnten, weil erst dann zwei Testtage zur Verfügung standen. Auch starke Konstanzannahmen können nicht darüber hinwegtäuschen, dass ein anspruchsvoller und anregender Deutsch- und Englischunterricht sowohl sprachliche Kompetenzen als auch kognitive Grundfähigkeiten beeinflussen kann. Wenn man bei DESI den Einfluss der kognitiven Grundfähigkeiten statistisch kontrolliert, geht man daher konservativ vor, d.h. man riskiert eher, den Erklärungswert von Unterrichtsmerkmalen zu verkennen, als ihn zu überschätzen.

In Abwägung all dieser Argumente sprechen durchaus gute Gründe dafür, unter Umständen – **je nach Fragestellung – ein eher sparsames Hintergrundmodell zu verwenden**⁴.

Als Fazit und Leitlinie lässt sich festhalten, dass Unterrichtsanalysen in DESI theoretisch und methodisch so erfolgen, dass (a) eine Vielfalt von Unterrichtszielen und -dimensionen einbezogen werden können, (b) die Wechselwirkung des Geschehens auf Klassenebene mit individuellen Prozessen berücksichtigt wird, und (c) Einflussgrößen außerhalb des Unterrichts (z.B. kognitive Grundfähigkeiten) statis-

3 Die Modelle konvergieren dann nicht, so dass z.B. das Merkmal „Kognitive Grundfähigkeit“ aus dem Modell eliminiert werden muss.

4 In diese Richtung zielte auch der Diskussionsbeitrag von Petra Stanat beim DESI-Symposium der DGfE, Frankfurt, März 2006.

tisch kontrolliert („herausgerechnet“) werden. Bei einer solchen Vorgehensweise arbeitet die Forschung eher vorsichtig, riskiert eher eine Über-Korrektur ihrer Befunde als das vorschnelle Identifizieren von vermeintlich wirksamen Faktoren. Umso bemerkenswerter ist es, dass DESI – wie die folgenden Kapitel belegen – sehr klare Aussagen darüber macht, was guter Unterricht in Englisch und Deutsch ist.

Literatur

- Baumert, J./Kunter, M./Brunner, M./Krauss, S./Blum, W./Neubrand, M. (2004): Schule und Unterricht. Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In: Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M./Pekrun, R./Rolf, H.-G./Rost, J./Schiefele, U. (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann, S. 314-349.
- Baumert, J./Trautwein, U./Artelt, C. (2003): Schulumwelten - institutionelle Bedingungen des Lehrens und Lernens. In: Deutsches PISA - Konsortium (Hrsg.): PISA 2000 - Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen: Leske + Budrich, S. 261-332.
- Berliner, D., C. (1995): Teacher expertise. In: L. W. Anderson (Hrsg.): International Encyclopedia of Teaching and Teacher Education. Oxford: Pergamon, S. 46-52.
- Blömeke, S. (2004): Empirische Befunde zur Wirksamkeit der Lehrerbildung. In: Blömeke S./Reinhold, P./Tulodziecki, G./ Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Braunschweig: Klinkhardt, S. 59-91.
- Blömeke, S. (2006): Voraussetzungen beim Lehrer. In: Arnold, K.-H./ Sandfuchs, U./Wiechmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt, S. 162-167.
- Bromme, R. (1997): Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In: Weinert, F. E. (Hrsg.): Psychologie des Unterrichts und der Schule 3. Göttingen: Hogrefe, S. 177-212.
- Brophy, J. E. (2000): Teaching (Vol. 1). Brussels: International Academy of Education & International Bureau of Education. URL: <http://www.ibe.unesco.org>.
- Brunner, M./Kunter, M./Krauss, S./Klusmann, U./Baumert, J./Blum, W./Neubrand, M./Dubberke, T./Jordan, A./Löwen, K./Tsai, Y.-M. (2006): Die professionelle Kompetenz von Mathematiklehrkräften: Konzeptualisierung, Erfassung und Bedeutung für den Unterricht. Eine Zwischenbilanz des COACTIV-Projekts. In: Prenzel, M./Allolio-Näcke, L. (Hrsg.): Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms. Münster: Waxmann, S. 54-82.
- Clausen, M. (2002): Unterrichtsqualität: Eine Frage der Perspektive? Münster: Waxmann.
- Ditton, H. (2006): Unterrichtsqualität. In: Arnold, K.-H./ Sandfuchs, U./Wiechmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt, S. 235-243.
- Fend, H. (1981): Theorie der Schule. München: Urban & Schwarzenberg.
- Gudjons, H. (2003): Frontalunterricht – neu entdeckt. Integration in offene Unterrichtsformen. Bad Heilbrunn: Klinkhardt.
- Haag, L./Lohrmann, K. (2006): Lehrerhandeln: Lehrerkognitionen und Lehrerexpertise. In: Arnold, K.-H./Sandfuchs, U./ Wiechmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbrunn: Klinkhardt, S. 617-626.
- Helmke, A. (2004): Unterrichtsqualität: Erfassen, Bewerten, Verbessern. Seelze: Klett-Kallmeyer.
- Helmke, A. (2007). „Einblick in die Lehr-Lern-Situation“: Ein Bogen zur Unterrichtsbeobachtung. In: Riecke-Baulecke, T. (Hrsg.): SchulleitungPlus: Grundsätze und Verfahren wirksamer Führung. München: Oldenbourg, S. 88-101/208-209.

- Helmke, A./Schrader, F.-W. (in Druck): Qualitätsmerkmale „guten Unterrichts“. In: Hellekamps, S./Plöger, W./Wittenbruch, W. (Hrsg.): Handbuch Erziehungswissenschaften, Band 4: Schule. Paderborn: Schöningh.
- Klieme, E. (2002): Was ist guter Unterricht? Ergebnisse der TIMSS-Video studie im Fach Mathematik. In: Bergsdorf, W./Court, J./Eckert, M./Hoffmeister, H. (Hrsg.): Herausforderungen der Bildungsgesellschaft. 4. Ringvorlesung der Universität Erfurt. Weimar: Rhino Verlag, S. 89-113.
- Klieme, E./Rakoczy, K. (2003): Unterrichtsqualität aus Schülerperspektive: Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht. In: Deutsches PISA - Konsortium (Hrsg.): PISA 2000 - Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen: Leske + Budrich, S. 333-360.
- Klieme, E. (2006): Empirische Unterrichtsforschung: Aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einleitung in den Thementeil. In: Zeitschrift für Pädagogik 52, H. 6, S. 765-773.
- Klieme, E./Lipowsky, F./Rakoczy, K./Ratzka, N. (2006): Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In: Prenzel, M./Allolio-Näcke, L. (Hrsg.): Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms, S. 128-146.
- Lanahan, L./McGrath, D. J./McLaughlin, M./Burian-Fitzgerald, M./Salganik, L. (2005): Fundamental Problems in the Measurement of Instructional Processes: Estimating Reasonable Effect Sizes and Conceptualizing What is Important to Measure. Washington: American Institutes for Research.
- Lipowsky, F. (2006): Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In: Zeitschrift für Pädagogik 51 (Beiheft "Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf"), S. 47-70.
- Meyer, H. (2004): Was ist guter Unterricht? Berlin: Cornelsen.
- Muthen, L. K./Muthen, B. O. (2003): Mplus User's Guide. Los Angeles: Muthen & Muthen.
- Pauli, Ch./Reusser, K. (2006): Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. In: Zeitschrift für Pädagogik 52, H. 6, S. 774-798.
- Shuell, T. J. (1996): Teaching and learning in a classroom context. In: Berliner, D. C./Calfee, R. C. (Hrsg.): Handbook of Educational Psychology. New York: Macmillan, S. 726-764.
- Weinert, F. E. (1992): Wie groß ist der Einfluß der Schule auf die geistige Entwicklung der Schüler, - und wie groß könnte der sein? Sonderreihe Öffentliche Vorträge 7, S. 5-22.
- Weinert, F. E./Helmke, A. (Hrsg.) (1997): Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union.
- Wellenreuther, M. (2005): Lehren und Lernen - aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht. Baltmannsweiler: Schneider Verlag Hohengehren.

Holger Ehlers / Nina Jude / Eckhard Klieme / Andreas Helmke /
Wolfgang Eichler / Heiner Willenberg

27 Soziodemografische und fachdidaktisch relevante Merkmale von Deutsch-Lehrpersonen

In diesem Kapitel geht es – anknüpfend an das im Kapitel 26 dargestellte Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts – um Merkmale von Lehrpersonen, die für das unterrichtliche Handeln von Bedeutung sein können. Sie spielen in der bildungspolitischen Diskussion eine wichtige Rolle, weniger jedoch in der fachdidaktischen Forschung. Kunze hat gezeigt, dass „Deutschlehrer und -lehrerinnen sowie deren berufliches Wissen und deren berufliche Kompetenzen (...) innerhalb der Fachdidaktik kein wichtiger Gegenstand systematischer empirischer Forschungen“ sind, und sie spitzt zu: „Sie sind derzeit eine Marginalie“ (Kunze 2004, S.191). DESI bietet erstmals die Chance, ausgewählte soziodemografische und berufsbiographische Merkmale von Lehrkräften wie Ausbildung, Fortbildung, Berufserfahrung sowie Aspekte der fachdidaktischen und pädagogischen Expertise anhand einer bundesweit repräsentativen Stichprobe zu untersuchen. Es muss jedoch darauf hingewiesen werden, dass wir uns aus Platzgründen nur auf eine Auswahl der bei DESI insgesamt erfassten Lehrermerkmale beschränken können. In diesem Kapitel geht es ausschließlich um die *Beschreibung* der genannten Personenmerkmale.

Zusammenhänge zwischen Lehrermerkmalen und dem Leistungsniveau der Schülerinnen und Schüler bleiben weiterführenden Analysen überlassen, ebenso Vergleiche zwischen verschiedenen Fächern. Ergebnisse zu den Merkmalen von Englisch-Lehrpersonen finden sich im Kapitel 30.

27.1 Stichprobenbeschreibung

An der Befragung der Deutsch-Lehrpersonen nahmen zu Beginn des Schuljahres 2003/04 358, am Ende 332 Personen teil. Legt man die Beteiligung am Ende des Schuljahres zugrunde, dann lässt sich die Zusammensetzung der Lehrer-Stichprobe (in Prozent) wie folgt beschreiben:

- *Schulart*: Hauptschule (72 Klassen, 22%), Realschule (102 Klassen, 31%), Gymnasium (110 Klassen, 33%), Integrierte Gesamtschule (24 Klassen, 7%), Schulen mit mehreren Bildungsgängen (24 Klassen, 7%).
- *Geschlecht*: 65% Frauen, 35% Männer.
- *Alter*: bis 31 J. (8%), 31-35 J. (12%), 36-40 J. (13%), 41-45 J. (13%), 46-50 J. (19%), 51-55 J. (18%), 56-60 J. (13%) und über 60 J. (4%).
- *Beschäftigungsstatus*: Vollzeit (59%), Teilzeit (41%).

Bei den im Folgenden berichteten Ergebnissen handelt es sich stets um gewichtete Daten. Die Befunde beziehen sich auf Lehrkräfte, die Deutschunterricht in der neunten Jahrgangsstufe geben. Prozentwerte geben jeweils den Prozentanteil von Schülern an, die von entsprechenden Lehrpersonen unterrichtet werden (vgl. auch Kapitel 3).

DESI kann auch in diesem Kapitel keine Angaben zu länderspezifischen Verhältnissen machen. Dass entsprechende Unterschiede durchaus relevant sein können zeigt eine Untersuchung zum Lese- und Medienverhalten von Schülerinnen und Schülern zur Lesesozialisation im Fach Deutsch. Gattermayer (2003) analysierte 1999 eine ähnlich große Stichprobe ($N = 359$) von Deutschlehrkräften aller allgemeinbildenden Schulen – allerdings beschränkt auf die Länder Bayern und Sachsen. Am Beispiel der Variable *Geschlecht* werden dort sehr große Unterschiede zwischen den Teilsamples aus Sachsen und Bayern berichtet, die „unweigerlich Auswirkungen auf den Vermittlungsprozess und mithin den Lese- und Literaturunterricht haben“ (ebd., S. 51). Aus dieser bisher nicht empirisch erhärteten Vermutung ergibt sich die Forderung, bei vertiefenden Analysen auch der DESI-Daten das Geschlecht der Lehrperson zu kontrollieren.

27.2 Ausbildung

Ein wichtiges Thema der bildungspolitischen und fachdidaktischen Debatten seit Jahrzehnten ist die professionelle Ausbildung der Lehrerschaft. Einen Einblick in die deutschsprachige Diskussion der neunziger Jahre geben z.B. die Sammelbände „Literaturstudium und Deutschunterricht auf neuen Wegen“ (Lecke 1996) und „Literatur und Medien in Studium und Deutschunterricht“ (Lecke 1999). Auch in aktuellen nationalen und internationalen Beiträgen wird diese Frage anhaltend und intensiv diskutiert (Berliner 2006; Lipowsky 2006). Die derzeitigen Entwicklungen, die auch stark mit der Umstellung auf das Bachelor-Master-Modell verbunden sind, führen zu Neubestimmungen und werden wiederum Korrekturen erfahren. Für DESI haben wir – vor dieser Reform – Daten erhoben, die grundsätzlicher Art sind und die formale Seite von Studiengängen bzw. -abschlüssen in das Blickfeld rücken.

Welche Rolle spielt die Art der Ausbildung für die Tätigkeit als Deutschlehrkraft?

Hier wurde zweistufig differenziert:

- ob dieses Fach studiert wurde oder nicht
- falls ja: ob Deutsch als Haupt- oder Erweiterungsfach studiert wurde; falls nein: welche Art der Qualifikation dann zugrunde liegt.

Eine Betrachtung über alle Schularten ergibt, dass 90% der Schülerinnen und Schüler im Fach Deutsch von Lehrkräften unterrichtet werden, die dieses Fach im Rahmen ihres Lehramtsstudiums oder eines anderen Studiengangs studiert haben, davon in den meisten Fällen (83%) als Hauptfach. Von den Lehrkräften, die Deutsch *nicht* als Studienfach belegt hatten, gibt etwa jede zweite an, durch einen „Lehrgang“ die Qualifikation erworben zu haben; jede dritte Lehrkraft wählte die Antwortmöglichkeit „Nachqualifikation“ und jede zehnte „Schnellkurs“ (vgl. Tabelle 27.1).

Tabelle 27.1: Ausbildung nach Schulart (Angaben in Prozent).

Schulart	Deutschstudium: Ja		Deutschstudium: Nein		
HS	69		31		
RS	94		6		
GY	100		0		
IGS	91		9		
MBG	95		5		
	Hauptfach	Nebenfach	Nachqualifi- kation	Schnellkurs	Lehrgang
HS	39	61	25	18	57
RS	90	10	23	0	77
GY	97	3	0	0	0
IGS	68	31	45	0	55
MBG	100	0	100	0	0

Schulartunterschiede bezüglich der Frage, ob Deutsch überhaupt studiert wurde, zeigen sich erwartungsgemäß vor allem zwischen Gymnasium (Deutschstudium bei 100%) und Hauptschule (69%). Dass eine hundertprozentige Nachqualifikation an den Schulen mit mehreren Bildungsgängen (MBG) stattfand, erklärt sich daher, dass diese Schulen in der DESI-Stichprobe nur aus einigen neuen Bundesländern stammen und somit die Umbruchsituation nach dem Beitritt abgebildet wird.

27.3 Fortbildung

Wie häufig bilden sich Deutsch-Lehrkräfte fort? Die Frage bezieht sich hier nicht auf jedwede Fortbildungen, sondern nur auf solche, die unmittelbar mit dem Deutschunterricht zu tun haben, wie z.B. mit fachdidaktischen Fragestellungen oder Unterrichtsmethoden. Wir fragten die Lehrkräfte, an wie vielen Tagen sie in den vergangenen zwei Jahren an entsprechenden Veranstaltungen teilgenommen haben. Betrachtet man alle Schularten zusammen, so nutzten zwei Drittel der Lehrkräfte Angebote der Lehrerfortbildung (siehe Tabelle 27.2).

Tabelle 27.2: Fortbildung für den Deutschunterricht in den letzten zwei Schuljahren nach Schulart (in Prozent).

Fortbildung	HS	RS	GY	IGS	MBG	Gesamt
Keine	44	39	28	17	12	31
1 Tag	6	7	9	21	6	8
2 Tage	11	26	22	37	28	24
3-5 Tage	28	16	23	20	31	23
6 Tage und mehr	11	12	18	5	23	14

Deutschlehrkräfte an Hauptschulen besuchten mit 44% am häufigsten keine fachspezifischen Fortbildungen. Aber auch an Realschulen und Gymnasien bleibt ein großer Teil der Lehrkräfte (39% und 28%) Fortbildungen, die sich unmittelbar auf den Unterricht im Fach Deutsch beziehen, fern. An den Gesamtschulen und den Schulen

mit mehreren Bildungsgängen nahmen prozentual deutlich mehr Lehrkräfte an fachspezifischen Fortbildungen teil (Nichtteilnahme bei 17% und 12%).

27.4 Expertise und Kooperation der Lehrkräfte

Pädagogische Psychologie und Unterrichtsforschung gehen davon aus, dass die Expertise von Lehrkräften, d.h. Umfang, Qualität, Aktualität und Vernetztheit ihres unterrichtsrelevanten Wissens eine wichtige Voraussetzung für das Gelingen des Unterrichts ist (Bromme 1997). Der Aufbau solcher für das praktische Handeln nutzbarer Wissenssysteme dürfte entscheidend von der Art und dem Umfang der Erfahrungen abhängen, die im Zuge der Auseinandersetzung mit beruflichen Aufgabenstellungen erworben werden. Insofern sind aus unserer Sicht sowohl die Beschäftigung mit unterrichtsrelevanten Themen als auch spezielle Funktionen im Rahmen des Kollegiums wichtige Quellen der Expertise. Bei DESI wurde Expertise aus den Bereichen Fachdidaktik und Pädagogik berücksichtigt:

- *Fachdidaktische Expertise*: Hier wurde u.a. nach dem Bezug einer fachdidaktischen Zeitschrift (Einzelitem) gefragt. Man kann davon ausgehen, dass eine Lehrkraft, die dieses Instrument regelmäßig nutzt oder es wenigstens ab und zu als Anregung ihrer Lehrtätigkeit versteht, sich professionelles Wissen aneignet und damit möglicherweise über günstigere Voraussetzungen für einen fachdidaktisch anspruchsvollen Unterricht verfügt.
- *Pädagogische Expertise*: Hier gehen wir davon aus, dass die Übernahme von Funktionen in der Ausbildung von Referendaren (Seminarleiter, Mentor) und Praktikanten als Beleg für die pädagogische Kompetenz gelten kann.

Weiterhin wird in der Schulforschung angenommen, dass Kooperation im Kollegium die Qualität der pädagogischen Arbeit verbessern kann. Die Ergebnisse zu den einzelnen Aspekten der Lehrerexpertise sowie zur Kooperationspraxis stellen sich folgendermaßen dar:

Eine fachdidaktische oder eine andere auf Fragen des Unterrichtens bezogene Zeitschrift wird – über alle Schularten betrachtet – von jeder zweiten Lehrkraft (46%) abonniert. Das größte Interesse zeigen dabei Realschullehrerinnen und -lehrer (56%), gefolgt von Lehrkräften an Hauptschulen (43%) und an Gymnasien (43%). Deutlich geringer sind die Anteile bei den Lehrkräften an Integrierten Gesamtschulen (33%) sowie an den Schulen mit mehreren Bildungsgängen (27%).

Als wichtige Indikatoren für das berufliche Engagement und die pädagogische Expertise betrachten wir die Wahrnehmung von Funktionen, wie z.B. Vorsitz in der Fachkonferenz, Betreuung von Praktikanten und Referendaren sowie andere Formen der Beteiligung an freiwilligen und/oder zusätzlichen Angeboten im Zusammenhang mit der Tätigkeit als Deutschlehrkraft. Unsere Fragen bezogen sich dabei auf das laufende Schuljahr.

Jede zehnte befragte DESI-Deutschlehrkraft (gewichtet 12%) arbeitet als Fachkonferenzleiter/in, Fachleiter/in oder Fachbetreuer/in für das Fach Deutsch. Deutlich höher ist – vor allem in Gymnasien – der Anteil jener, die Praktikanten/innen (z.B.

Studierende im Blockpraktikum) oder Referendare bzw. Lehramtsanwärter betreuen (gewichtet: 30%).

Von einem Engagement bei freiwilligen schulinternen Angeboten im Bereich Deutsch (Leitung einer Deutsch Arbeitsgemeinschaft, Projekte etc.) berichtet insgesamt jede sechste Deutschlehrkraft. Solche Angebote finden im Gymnasium eher statt als an Hauptschulen (Anteil der Beteiligten: 21% bzw. 11%). Integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen wiederum haben einen Schwerpunkt bei außerunterrichtlichen Förderkursen (mitwirkend hier: jeweils 15%, an den übrigen Schularten unter 10%). Ein interessanter Unterschied zwischen diesen beiden Schularten liegt in der Häufigkeit, mit der Lehrkräfte Nachhilfe im Fach Deutsch geben: In den (ostdeutschen) Schulen mit mehreren Bildungsgängen sind daran 28% beteiligt, in den (überwiegend westdeutschen) Integrierten Gesamtschulen nur 7%, ansonsten immerhin um 16%. Offenbar kommt diese Art des Zusatzunterrichts nicht sehr selten vor.

Die „Zusammenarbeit mit anderen Lehrkräften der eigenen Schule im Fach Deutsch (z.B. gemeinsame Unterrichtsvorbereitung und/oder –durchführung; gemeinsame Erstellung von Unterrichtsmaterialien)“ wird von jeder zweiten Lehrkraft angegeben und stellt somit den am häufigsten auftretenden Indikator für berufliches Engagement dar, wobei Gymnasial- und Realschullehrer (jeweils unter 50%) weniger in Kooperationen eingebunden sind als die Lehrkräfte anderer Schularten (jeweils mehr als zwei Drittel). Von einer intensiveren, konzeptuell ausgerichteten Kooperation im Sinne der „Teilnahme an Arbeitsgruppen zur Verbesserung des Deutschunterrichts (auch in Kooperation mit anderen Schulen)“ berichtet insgesamt nur jede sechste Lehrkraft; auch hierbei zeigt sich das besonders hohe Kooperationsniveau an Integrierten Gesamtschulen.

Die berichteten schulartspezifischen Daten müssen sehr vorsichtig interpretiert werden, insbesondere wenn Tendenzen im Sinne von schularttypischen Kulturen sichtbar werden, die es weiter zu untersuchen gilt. Auf die Kooperation im Fachkollegium geht Kapitel 34 näher ein, wo auch Zusammenhänge mit der Leistung der Schülerinnen und Schüler diskutiert werden.

Literatur

- Berliner, D. C. (2006): The dangers of some new pathways to teacher certification. In: Oser, F. K./Achtenhagen, F./Renold, U. (Hrsg.): *Competence oriented teacher training. Old research demands and new pathways*. Rotterdam: Sense Publishers, S. 117-127.
- Bromme, R. (1997): Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In: Weinert, F. E. (Hrsg.): *Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Pädagogische Psychologie, Vol. 3)*. Göttingen: Hogrefe, S. 177-212.
- Ditton, H./Merz, D./Edelhäuser, T. (2002): Einstellungen von Lehrkräften und Schulleiter/innen zu zentralen Testuntersuchungen an Schulen. In: *Empirische Pädagogik* 16, H.1, S. 17-33.
- Gattermaier, K. (2003): *Literaturunterricht und Lesesozialisation*. Regensburg: edition vulpes.
- Kunze, I. (2004): *Konzepte von Deutschunterricht*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lecke, B. (1996): *Literaturstudium und Deutschunterricht auf neuen Wegen*. Frankfurt: Lang.

- Lecke, B. (1999): Literatur und Medien in Studium und Deutschunterricht. Frankfurt: Lang.
- Lipowsky, F. (2006): Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In: Zeitschrift für Pädagogik 51. Beiheft „Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf“, S. 47-70.

Eckhard Klieme / Nina Jude / Dominique Rauch / Holger Ehlers /
 Andreas Helmke / Wolfgang Eichler / Günther Thomé / Heiner Willenberg

28 Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts

Nicht nur in Deutschland oder im deutschsprachigen Raum, sondern auch im internationalen Diskurs lassen sich unterschiedliche Konzepte des schulischen Unterrichts für die Muttersprache (besser: Verkehrssprache) unterscheiden. Sawyer und Van den Ven (2007) identifizieren vier Paradigmen, mit denen sie die Veränderung besonders der Sprachlehre im muttersprachlichen Unterricht seit dem 19. Jahrhundert beschreiben: Ausgangspunkt war das „akademische“ Paradigma, bei dem die Vermittlung von grammatikalischem Wissen im Vordergrund stand. Zu Beginn des 20. Jahrhunderts wurde dies abgelöst durch ein „entwicklungsorientiertes“ Paradigma, das sprachpsychologische Befunde aufgriff. Normativer Grammatikunterricht wurde jetzt ersetzt durch die Unterstützung der kindlichen Sprachentwicklung. Den 1960er und 70er Jahren ordnen die Autoren ein „kommunikatives“ Paradigma zu, bei dem – häufig in emanzipatorischer Absicht – Einsichten in die soziale Funktion von Sprache vermittelt werden sollten. Als neueste Entwicklung schließlich bezeichnen sie ein „utilitaristisches“ Paradigma, das sich u. a. auf psychometrische Studien zu sprachlichen Fähigkeiten und zu Effekten des Sprachunterrichts stützt. Auch hier stehen kommunikative Aspekte im Vordergrund, allerdings nicht mehr so sehr aus einer sozialkritischen Perspektive, sondern im Blick auf alltägliche Transaktionen. In diesem Unterricht soll die *alltägliche Kommunikationsfähigkeit* gefördert werden. Insgesamt charakterisieren die Autoren die konzeptionelle Entwicklung des muttersprachlichen Unterrichts als „a shift from teaching grammar to developing communicative competence“ (Sawyer/Van den Ven 2007, S. 6).

In Deutschland galt lange nicht die Sprachlehre sondern das Lesen und Verstehen von Literatur als Königsdisziplin des Deutschunterrichts. Bis in die siebziger Jahre des zwanzigsten Jahrhunderts war die Fachdidaktik bestrebt, die germanistischen Kategorien, wenn auch vereinfacht, in den Unterricht einzubringen. Dieser teilweise unpädagogischen Dominanz der Fachsystematik stellte sich die Rezeptionsforschung entgegen, die auf die tatsächlichen Verstehensfähigkeiten der Kinder und Jugendlichen abhob, so dass ab den achtziger Jahren der Boden für einen kreativen und produktiven Unterricht bereitet war.

Neuere Arbeiten zum Deutschunterricht bestätigen die Tendenz zu einer „utilitaristischen“ Orientierung. Überlegungen zu einem Leseunterricht, in dem Gebrauchstexte berücksichtigt werden, und eine explizite Schreibdidaktik werden in der Fachdiskussion zunehmend wichtiger (vgl. etwa Kämper-van den Boogaart 2003). In Auseinandersetzung mit Konzepten sprachlicher Kompetenz, wie sie auch die DESI-Studie nutzt, wird neuerdings ein „kompetenzorientierter Unterricht“ propagiert, der dem Schüler Anregungen und Freiräume zur Entwicklung sprachlicher

Kompetenzen, insbesondere im Bereich des Schreibens und des Lesens, sichern soll (Dehn/Hüttis-Graff 2005).

Dabei spielt v.a. die Auseinandersetzung mit Ergebnissen von PISA 2000 (Sekundarstufe I) aber auch mit IGLU (Grundschule) eine wichtige Rolle; sie geben Anlass zu einer grundlegenden Reflexion über Ziele und Erfolge des sprachlichen und literarischen Unterrichts (vgl. Abraham/Bremerich-Vos/Frederking/Wieler 2003).

Auch dort, wo Kompetenzkonzepte der Schulleistungsstudien aufgegriffen werden, bleibt die fachdidaktische Diskussion jedoch weitgehend normativ und praxeologisch. Man arbeitet teils bewährte, teils innovative Unterrichtskonzepte auf in der Erwartung, kommunikative, anwendungsrelevante sprachliche Kompetenzen, insbesondere Schreib- und Lesekompetenz, systematisch fördern zu können. Ein typisches Beispiel hierfür ist etwa das von Eikenbusch (2001) verfasste Buch „Qualität im Deutschunterricht der Sekundarstufe I und II“. Der Autor greift die neuere Qualitätsdebatte auf und gibt umfangreiche Empfehlungen zur „Konfiguration“ des Deutschunterrichts. Dass die vorgeschlagenen Konzepte erfolgreich sind, kann jedoch an keiner Stelle empirisch belegt werden. Dies ist nicht dem Autor als Mangel zuzuschreiben, sondern dem Erkenntnisstand der Fachdidaktik bzw. der auf den Deutschunterricht ausgerichteten Lehr-Lern-Forschung. Vor der DESI-Studie wurden so gut wie keine breiten, repräsentativen Erhebungen zur Praxis des Deutschunterrichts durchgeführt, ganz zu schweigen von systematischen, experimentellen oder quasi-experimentellen Wirkungsanalysen, mit denen Effekte bestimmter Unterrichtsmerkmale hätten nachgewiesen werden können. „Die Praxis des Deutschunterrichts der letzten Jahrzehnte ist nur spärlich dokumentiert. Zugänglich sind allenfalls didaktische und methodische Programmschriften. Auf die Frage, inwiefern die faktische Praxis diesen Programmen entsprach, kann eine verlässliche Antwort nicht gegeben werden.“ (Bremerich-Vos 2006, S. 450; gleichlautend Eikenbusch 2001, S. 52).

Wie können fachdidaktisch reflektierte, empirisch gestützte Aussagen über „guten Deutschunterricht“ – gemessen an jeweils zu spezifizierenden Erfolgskriterien – gemacht werden, die über Praxeologie und Präskription hinausgehen? Ein erster Ansatz kann sicherlich darin bestehen, aus den spezifischen Leistungsdefiziten von Schülerinnen und Schülern, wie sie etwa in neueren Schulleistungsstudien offenbar geworden sind, gleichsam remediale Maßnahmen abzuleiten (vgl. etwa Abraham et. al. 2003). Entsprechende Vorschläge finden sich auch im vorliegenden Band in den Abschnitten zu Kompetenzbereichen im Deutschen. Willenberg (2007) und seine Mitautoren haben in diesem Sinne eine Vielfalt von Maßnahmen für den Deutschunterricht vorgeschlagen, die mit Befunden der DESI-Studie begründet werden und darauf abzielen, dass sich der Unterricht an den beschriebenen Kompetenzniveaus orientiert. Ein zweiter, direkt empirisch stringenterer Ansatz besteht darin, im Anschluss an Schulleistungsstudien – etwa unter Nutzung darin entwickelter Tests – spezielle Trainingsmaßnahmen zu evaluieren. Eine explorative

Studie, die in diesem Sinne ein Lesestrategietraining beschreibt und unter Nutzung von DESI-Aufgaben auswertet, hat Frey (2007) vorgestellt.

Im vorliegenden Kapitel wird ein anderer Ansatz verfolgt: Wir nutzen die DESI-Daten selbst, um systematisch auf den Unterricht zu schauen und Zusammenhänge zwischen Lernvoraussetzungen, Unterrichtsprozessen und Lernergebnissen zu ermitteln. Zwar kann DESI für den Bereich des Deutschen – anders als im Englischen (vgl. Kap. 29) – keine Videoanalysen vorlegen, aber immerhin war es möglich, sowohl Schülerinnen und Schüler als auch Lehrkräfte differenziert nach der Praxis des Deutschunterrichts, deren Wahrnehmung und Bewertung zu fragen. Auf dieser Basis lässt sich die Alltagspraxis des Deutschunterrichts in wichtigen Aspekten beschreiben und es können systematisch Zusammenhänge mit der Leistungsentwicklung im Verlauf der 9. Jahrgangsstufe untersucht werden. Die Analysen folgen somit dem üblichen Ansatz der Unterrichtsqualitätsforschung (vgl. Klieme 2006 sowie Kap. 26). Sie greifen bewusst die allgemeinen Beschreibungskategorien für Unterrichtsqualität auf, wie sie in dieser Forschung herausgearbeitet worden sind – bislang überwiegend am mathematischen und naturwissenschaftlichen Unterricht. Solche Qualitätsdimensionen wie effiziente Klassenführung, Klarheit und Konkretheit der Ziele und Aufgaben, Strukturiertheit, hohe Leistungserwartungen, adaptive Unterrichtsformen, fachliche und persönliche Unterstützung der Schüler wurden von Eikenbusch (2001, S. 85) als relevant auch für den Deutschunterricht postuliert, allerdings ohne empirische Belege. Ziel des vorliegenden Kapitels ist es, die Wirksamkeit dieser Grunddimensionen der Unterrichtsqualität im Bereich des Deutschunterrichts nachzuweisen, soweit dies im Rahmen von DESI – allein gestützt auf Fragebögen und Tests, ohne Unterrichtsbeobachtungen – möglich ist.

Hierbei gehen wir in drei Schritten vor: Im ersten und zweiten Schritt wird der Deutschunterricht aus der Sicht der Lehrpersonen (Abschnitt 28.1) bzw. der Schülerinnen und Schüler (Abschnitt 28.2) beschrieben. Wir stellen die hierzu verwendeten Fragebogenskalen sowie ausgewählte Detailbefunde vor, untersuchen mittels Faktorenanalysen die grundlegenden Dimensionen, in denen der Unterricht wahrgenommen wird, und fragen schließlich, ob sich je nach Bildungsgang spezifische Unterrichtsmilieus identifizieren lassen. Die wichtigsten analytischen Ergebnisse finden sich im Abschnitt 28.3: Hier werden die von den jeweils Beteiligten wahrgenommenen Unterrichtsmerkmale systematisch zur Kompetenzentwicklung der Schülerinnen und Schüler in Beziehung gesetzt. Zum einen soll untersucht werden, ob Deutschunterricht je nach Eingangsbedingungen der Klasse unterschiedlich gestaltet wird; zum anderen soll die „Wirksamkeit“ der Unterrichtsmerkmale im Sinne ihres Einflusses auf die Kompetenzentwicklung über ein Schuljahr hinweg beleuchtet werden. Rahmenkonstellationen und Eingangsbedingungen werden dabei ebenso wie individuelle Urteilsprozesse sorgfältig kontrolliert, um die Effekte der Unterrichtsmerkmale auf Klassenebene möglichst „rein“ abschätzen zu können.

28.1 Alltagspraxis des Deutschunterrichts aus Sicht von Lehrerinnen und Lehrern

In folgenden Abschnitt wird die Alltagspraxis des Deutschunterrichts aus Sicht von Deutschlehrkräften beschrieben. Die Darstellung beruht auf der Befragung von 370 Deutschlehrkräften. Folgende Fragen sollen sukzessive beantwortet werden: Wie viel Zeit haben die Lehrerinnen und Lehrer, das Fach Deutsch zu unterrichten? Welche Methoden – von der Organisationsform (z.B. Einzel- oder Gruppenarbeit) bis zum Einsatz von Lernerfolgskontrollen – wählen sie für ihren Unterricht? Welche Materialien und Textsorten werden verwendet? Und welche Bedeutung haben spezifische Maßnahmen zur Förderung sprachlicher Kompetenzen? Die Beschreibung der Unterrichtspraxis wird ergänzt um Angaben zur Kooperation zwischen den Lehrkräften und mündet schließlich in eine zusammenfassende Analyse von Grunddimensionen der Unterrichtsgestaltung, hinsichtlich derer die Bildungsgänge unterschiedliche Profile aufweisen.

Lehr-Lern-Zeit

Eine erste Annäherung an die Alltagspraxis des Deutschunterrichts stellt die rein quantitative Erfassung der zur Verfügung stehenden Lehr-Lern-Zeit dar. Das Fach Deutsch wird in der 9. Jahrgangsstufe zwischen 3 Stunden (26%) und 5 oder mehr Stunden (7%) pro Woche unterrichtet. Am weitaus häufigsten werden 4 Stunden pro Woche erteilt (67%).

In der Praxis wird nicht jede dieser Deutschstunden tatsächlich gehalten, deshalb wurden die Lehrkräfte am Ende der neunten Jahrgangsstufe nach dem Umfang des Unterrichtsausfalls im laufenden Schuljahr und der Art der Vertretung gefragt (siehe Tabelle 28.1).

Tabelle 28.1: Unterrichtsausfall und -vertretung im Fach Deutsch während der neunten Jahrgangsstufe.

Ausfall/Art der Vertretung	Mittelwert (Stunden)	Standardabweichung
Ersatzloser Ausfall	6.02	5.71
Deutsch-Vertretungslehrkraft	4.05	7.81
Fachfremde Vertretungslehrkraft	3.24	3.34

Insgesamt fallen durchschnittlich 13.31 Stunden im Schuljahr aus, was 3.5 Wochen ohne Deutschunterricht im Schuljahr entspricht. Die sehr hohen Standardabweichungen zeigen an, dass sich die Klassen hinsichtlich des Unterrichtsausfalls deutlich unterscheiden. Die große Streuung des Unterrichtsausfalls lässt sich am

Beispiel des ersatzlosen Ausfalls verdeutlichen: Während knapp 20% der Lehrkräfte¹ berichteten, dass ihr Unterricht im laufenden Schuljahr nie ersatzlos ausgefallen ist, geben 28% einen ersatzlosen Unterrichtsausfall von 10 Stunden oder mehr an; knapp 1% sprechen sogar von einem Unterrichtsausfall in der Größenordnung von 30 Stunden.

Findet der Unterricht statt, kann die effektive Lehr-Lern-Zeit durch organisatorische Bedingungen wie zum Beispiel den Wechsel des Unterrichtsraumes verkürzt werden. 94% der Lehrkräfte berichten, dass sie maximal 5 Minuten auf Raumwechsel verwenden müssen. Allerdings zeigen die 4.5% der Lehrkräfte, die angeben, dass die *durchschnittlich in jeder Unterrichtsstunde* durch Raumwechsel verloren gegangene Zeit 6 bis 10 Minuten beträgt, und diejenigen 1.4%, die sogar 11-15 Minuten angeben, dass auch dieser Faktor die zur Verfügung stehende Lehr-Lern-Zeit erheblich verkürzen kann. Gefragt nach dem gesamten Zeitverlust durch Vorgänge, „die nichts mit dem Unterrichtsstoff zu tun haben (z.B. Organisatorisches, Disziplinstörungen, Gespräche nichtfachlichen Inhalts)“, ergibt sich das in Abbildung 28.1 dargestellte Bild.

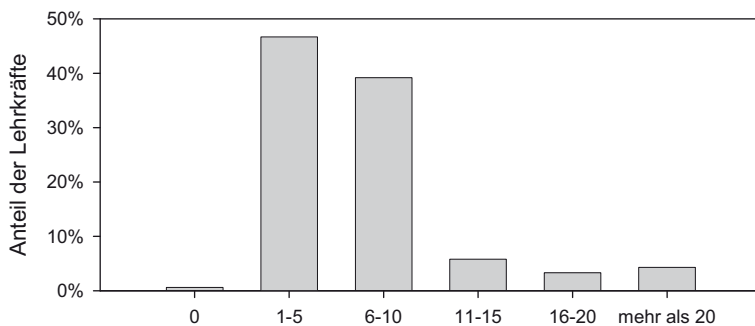


Abbildung 28.1: Zeitverlust durch unterrichtsfremde Vorgänge in Minuten.

Etwa 14% der Lehrkräfte verbringen nach eigenen Angaben pro Unterrichtsstunde mehr als 11 Minuten mit unterrichtsfremden Aspekten der Arbeit in der Klasse.

Unterrichtsmethoden

In der Deutschdidaktik wird eine angemessene Diversifikation der Methoden im Unterricht als wichtig und erstrebenswert erachtet (vgl. bspw. Fritzsche 1994a, 1994b). Empirische Forschungen in anderen Fächern sprechen allerdings in Übereinstimmung mit der Allgemeinen Didaktik dafür, dass Unterrichtserfolg weniger von den Methoden an sich als von der Qualität ihres Einsatzes abhängt (Klieme,

¹ Hier und im weiteren werden gewichtete Ergebnisse berichtet (vgl. Kap. 2); die Ergebnisse können auf die *Schüler* der neunten Jahrgangsstufe verallgemeinert werden, nicht auf die Lehrer. Das bedeutet, dass der obige Befund genau genommen lauten müsste „20% der Schüler werden von Lehrkräften unterrichtet, die berichteten, dass ihr Unterricht im laufenden Schuljahr nie ersatzlos ausgefallen ist“ (vgl. Kap. 27 und 30). Der Lesbarkeit halber werden im weiteren dennoch einfachere Formulierungen mit Bezug auf die Lehrerschaft verwendet.

2006). DESI bietet eine Gelegenheit, diese Thesen im Blick auf die Praxis des Deutschunterrichts zu überprüfen. Dies soll anhand ausgewählter Merkmale, nämlich der Organisationsformen des Unterrichts, der Differenzierung, der Partizipation sowie des Einsatzes von Hausaufgaben und Lernerfolgskontrollen erfolgen.

Organisationsformen des Unterrichts

Die Lehrkräfte der DESI-Klassen wurden v.a. nach der Verwendung verschiedener, überwiegend reformpädagogisch inspirierter Unterrichtsformen gefragt. Auf der Basis ihrer Antworten stellt Tabelle 28.2 dar, wieviel Prozent der Schülerinnen und Schüler regelmäßig (d.h. mindestens *ein paar Mal pro Monat*) diese Formen der Unterrichtsorganisation erfahren.

Tabelle 28.2. Organisationsformen des Deutschunterrichts nach Bildungsgang, Unterrichtsform wird mindestens mehrmals pro Monat realisiert, Häufigkeiten in Prozent.

Organisationsform	HS	RS	IGS	GY	gesamt
Arbeit mit kleinen Schülergruppen	64	40	67	60	54
Diskussionsrunden	61	45	48	57	53
Fachübergreifendes und fächerverbindendes Lernen	68	31	22	22	37
Freiarbeit	32	10	27	8	17
Wochenplan	15	14	22	10	14
Projektlernen	12	10	12	6	10
Peer-Tutoring	13	4	-	5	6
Geschlechtshomogene Kleingruppen	-	4	-	7	3
Lernzirkel & Stationslernen	10	3	-	3	5
Gemeinsame Unterrichtsvorbereitung	16	5	21	1	8
Gemeinsame Unterrichtsdurchführung	4	1	5	-	2

Bezüglich der Unterrichtsorganisation zeigen sich deutliche Schwerpunkte über alle Bildungsgänge hinweg: Kleingruppenarbeit und Diskussionsrunden im Deutschunterricht werden in über 40% aller Klassen mehrmals pro Monat organisiert. Der Deutschunterricht in der Hauptschule zeichnet sich darüber hinaus durch häufiges fächerübergreifendes Lernen und Freiarbeit aus. Insgesamt werden in diesem Bildungsgang die Alternativen zum Frontalunterricht häufiger angewendet als in den drei anderen. Zur Charakterisierung der Alltagspraxis des Deutschunterrichts werden im weiteren zwei Skalen verwendet, eine zur Kleingruppenarbeit und eine zu alternativen Unterrichtsmethoden. Letztere erfasst, wie häufig Methoden wie fächerübergreifendes und fächerverbindendes Lernen, Freiarbeit, Wochenplan und Lernzirkel/Stationslernen zum Einsatz kommen.

Differenzierung

Welche Formen der inneren Differenzierung wenden Deutschlehrerinnen und -lehrer an, um Schülerinnen und Schüler gezielt zu fördern? Die in DESI befragten Lehrkräfte gaben Auskunft auf einer vierstufigen Antwortskala (*trifft gar nicht zu*

bis *trifft genau zu*), ob sie in ihrer Unterrichtspraxis verschiedene Möglichkeiten des differenzierten Unterrichtens verwenden. Die Skala Differenzierung ($\alpha = .70$) umfasst neun Items, die verschiedene Formen der Differenzierung im Unterricht abbilden. Als Beispielitem sei hier die Differenzierung bei der Stillarbeit angegeben: „Bei der Stillarbeit variere ich die Aufgabenstellungen, um Schülern/innen unterschiedlicher Leistungsstärke gerecht zu werden“. Immerhin 46% der Deutschlehrkräfte vergeben bei der Stillarbeit verschiedene Aufgaben (*trifft eher zu* und *trifft genau zu* zusammengefasst), um dem unterschiedlichen Niveau ihrer Schüler gerecht zu werden. Die Praxis bei der Bildung von Kleingruppen sieht dagegen anders aus: Die Gruppenzusammenstellung leistungshomogen zu gestalten ist eher unüblich; nur 16% der Deutschlehrkräfte greifen auf dieses Mittel zurück. Dagegen bildet ein Großteil der Deutschlehrerinnen und -lehrer (70%) bei Gruppenarbeit bevorzugt leistungsheterogene Gruppen, die zwar dazu dienen können, dass die erfolgreicherer Lerner die etwas langsameren anregen, nicht aber als Form der Differenzierung angesehen werden können.

Partizipation

Eine weitere Frage, die DESI zu beantworten versucht, ist, wie es um die Möglichkeiten der Schülerinnen und Schülern zur Partizipation an der Gestaltung des Deutschunterrichts bestellt ist. Mit einer Skala ($\alpha = .70$) von fünf Items wurde erfragt, wie oft Deutschlehrkräfte Schülerinnen und Schülern die Möglichkeit geben, Inhalte und Reihenfolge von Lerninhalten mitzubestimmen, wie häufig in Klassenarbeiten und Hausarbeiten Aufgaben zur Auswahl gestellt werden und wie oft von Schülern verwendete Aufgaben zum Einsatz kommen. Die Lehrkräfte hatten die Möglichkeit, ihren Unterricht hinsichtlich der Häufigkeit von Partizipationsmöglichkeiten der Schülerinnen und Schülern einzuschätzen. Als Beispielitem sei hier die *Auswahl* von Lerninhalten vorgegeben: „Wie oft tun Sie folgendes? Ich beteilige Schüler/innen an der Auswahl von Lerninhalten aus vorgegebenen Bereichen“. Knapp ein Prozent der Deutschlehrkräfte geben ihren Schülern wöchentlich die Möglichkeit, die Auswahl der Lerninhalte mitzubestimmen, 13% stellen Inhalte einmal im Monat zur Wahl.

Didaktischer Stellenwert von Hausaufgaben

Hausaufgaben gehören zur Alltagspraxis des Unterrichts in Kernfächern und sind einschlägiges Thema der neueren Bildungsforschung. So konnten Trautwein und Kollegen (2001) für den Mathematikunterricht in der 7. Jahrgangsstufe zeigen, dass regelmäßige Hausaufgaben einen positiven Einfluss auf die Mathematikleistung haben. Lipowsky u.a. (2004) wiesen darüber hinaus nach, dass auch die Behandlung der Hausaufgaben im Unterricht einen relevanten Einfluss auf die Leistungen von Schülern hat.

Um mehr über die Hausaufgabenpraxis im Deutschunterricht zu erfahren, wurden die Lehrkräfte in DESI mittels einer Skala ($\alpha = .80$) aus sechs Items zum didaktischen Stellenwert von Hausaufgaben in ihrem Unterricht befragt. Sie hatten dabei die Möglichkeit, auf einer vierstufigen Antwortskala (*nie*, *ein paar mal*

pro Jahr, ein paar mal pro Monat, ein paar mal pro Woche) anzugeben, wie oft sie Hausaufgaben wie erfragt einsetzen. Hausaufgaben sind routinemäßiger Bestandteil des Deutschunterrichts: 73% der Deutschlehrkräfte nutzen mehrmals pro Woche die Möglichkeit, Hausaufgaben zu vergeben. Hausaufgaben werden im allgemeinen nicht zur Auswahl gestellt (69% der Lehrkräfte tun dies höchstens ein paar mal pro Jahr). Viele Lehrkräfte (75%) überprüfen routinemäßig, ob die Hausaufgaben gemacht wurden. Da sich die Wertschätzung sprachbezogener Schülerkompetenzen durch den Lehrer als wichtig für die Leistung im Fach Deutsch herausstellt (vgl. 28.3), sei hier darauf hingewiesen, dass 69% der Lehrkräfte angeben, mindestens ein paar mal pro Monat die Hausaufgaben „im Hinblick auf sprachlichen Ausdruck“ zu korrigieren. Bei einem Drittel der Deutschlehrkräfte gehört dies nicht zur Alltagspraxis des Deutschunterrichts.

Häufigkeit von Lernerfolgskontrollen

In DESI wurde erfragt, wie häufig bestimmte Formen der Lernerfolgskontrolle durchgeführt wurden².

Tabelle 28.3: Häufigkeit von Lernerfolgskontrollen.

Wie häufig haben Sie im laufenden Schuljahr in Deutsch die folgenden Arten von Lernerfolgskontrollen eingesetzt?	Mittelwert Anzahl	Standard- abweichung
Klassen- oder Kursarbeiten, Schulaufgaben	6.0	4.3
Vergleichsarbeiten, d.h. klassen- bzw. lerngruppenübergreifend gleiche Klassenarbeiten	1.3	2.3
informelle Tests mit Multiple-Choice-Antworten	0.3	1.0
informelle Tests mit offenen Antworten	1.6	2.0
Portfolio des Lernfortschritts	0.5	1.5

Es überwiegen klassische Klassenarbeiten und Kursarbeiten gegenüber allen anderen Formen der Lernerfolgskontrollen. Inwieweit sich dieses Bild nach Einführung der Bildungsstandards geändert hat, kann DESI nicht zeigen.

Unterrichtsinhalte und Materialien

Welche Bedeutung hat der Umgang mit literarischen Texten und Inhalten für den Deutschunterricht? Wird kreatives Schreiben der Schülerinnen und Schüler im Deutschunterricht gefördert? Im Folgenden soll zunächst dargestellt werden, welche Unterrichtsmaterialien im Deutschunterricht verwendet werden, um dann die Frage nach der Bedeutung literarischer Texte für die Alltagspraxis des Deutschunterrichts zu klären. Anschließend werden der didaktische Stellenwert von Prosa und Poesie und die Bedeutung des kreativen Schreibens für die alltägliche Gestaltung des Deutschunterrichts dargestellt.

² Die DESI Erhebung fand vor Einführung der Bildungsstandards statt.

Textsorten und Material

Eine wesentliche Frage bei der Gestaltung von Deutschunterricht, aber auch bei seiner Reflexion ist die nach den im Unterricht verwendeten Textsorten und anderen Unterrichtsmaterialien (bspw. Lehrbüchern, Videos etc.). In DESI wurde nach der Verwendung so unterschiedlicher Textsorten wie Gedichten und Werbeanzeigen gefragt. Die Lehrerinnen und Lehrer hatten die Möglichkeit, die Häufigkeit der Verwendung der Materialien durch Auswahl von *nie*, *einmal pro Jahr*, *einmal pro Monat* und *einmal pro Woche* einzuschätzen. Abbildung 28.2 weist aus, welchen Stellenwert – ob in produktiver oder rezeptiver Verwendung – die einzelnen Textsorten bzw. Materialien einnehmen.

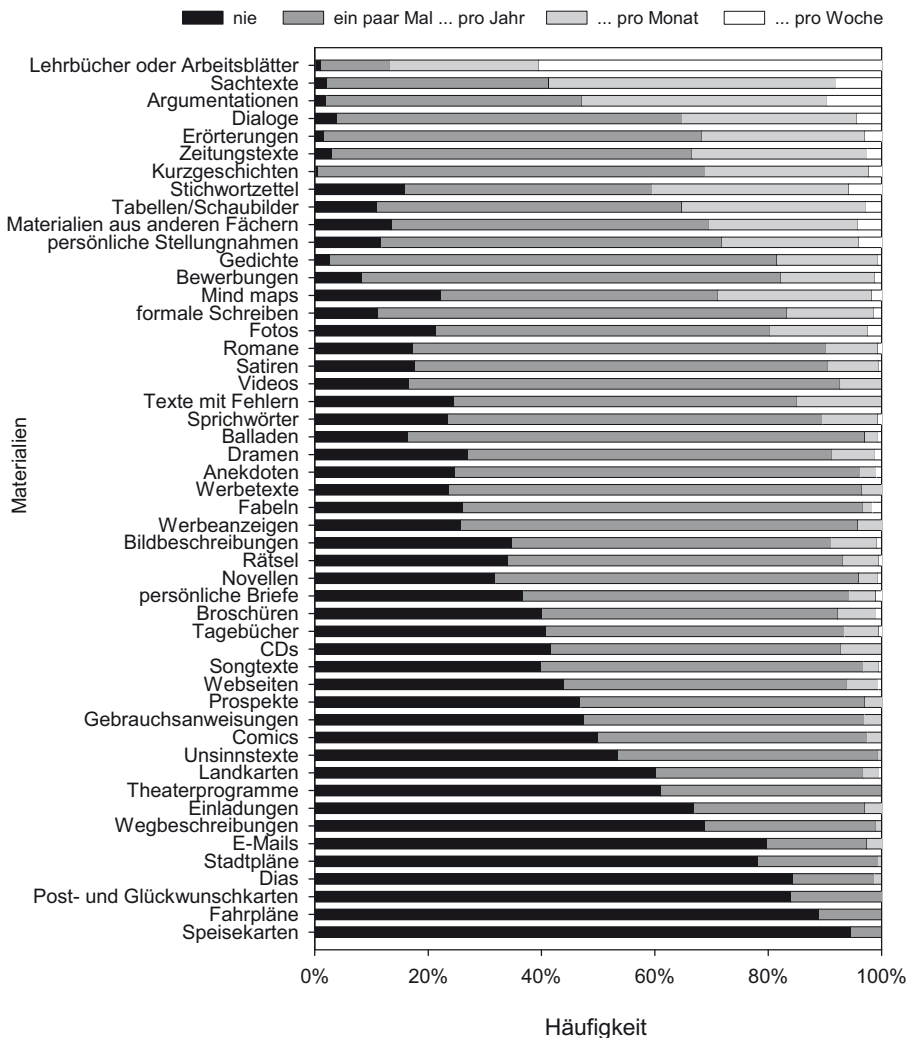


Abbildung 28.2 : Häufigkeit der Verwendung von Textsorten/Material (Itemebene).

Wie auch im Fach Englisch liegen Lehrbücher und Arbeitsblätter erwartungsgemäß deutlich an der Spitze. Ihnen folgen Sachtexte und Argumentationen. Dialoge, Erörterungen, Zeitungstexte und Stichwortzettel rangieren dann mit der ersten literarischen Textsorte, den Kurzgeschichten, auf etwa einem Niveau. So genannte „diskontinuierliche Textsorten“ wie z.B. Stadt- und Fahrpläne oder Speisekarten, die im PISA-Lesetest eine gewisse Bedeutung haben, führen im Deutschunterricht – wie übrigens auch in den Lehrplänen – ein Nischendasein. Durch das World Wide Web entstandene neue Textsorten, wie die in der Regel „kontinuierlichen“ E-Mails oder die in der Regel eher „diskontinuierlichen“ Webseiten, haben sich bereits als Unterrichtsgegenstände oder -medien etablieren können – wenngleich auf bescheidenem Niveau.

In der Diskussion der PISA-Ergebnisse wurde immer wieder die Vermutung vorgetragen, dass im Deutschunterricht literarische Texte zuungunsten insbesondere der Sachtexte überrepräsentiert sind. Um dies zu überprüfen, wurden die im Unterricht verwendeten Textsorten in drei übergeordneten Skalen zusammengefasst:

- Skala „Unterrichtsmaterialien: authentische Sachtexte“, umfasst z.B. Gebrauchsanweisungen, Werbeanzeigen, Internetseiten (28 Items, $\alpha = .87$, $M = 1.83$).
- Skala „Unterrichtsmaterialien: literarische Texte“, umfasst z.B. Anekdoten, Gedichte, Novellen (12 Items, $\alpha = .76$, $M = 1.86$).
- Skala „Unterrichtsmaterialien: Texte mit kommunikativem Bezug“ umfasst z.B. Bewerbungen, Einladungen, persönliche Briefe (8 Items, $\alpha = .61$, $M = 1.95$).

Es zeigt sich, dass die Lehrerschaft angibt, die Textsorten dieser drei Hauptkategorien gleich häufig zu benutzen. Insofern kann die Vermutung einer Unterrepräsentation von Sachtexten zurückgewiesen werden.

Unterrichtselemente Prosa

Mit Hilfe der Skala „Unterrichtselemente Prosa“ ($\alpha = .91$) aus 15 Items wurde erfragt, wie häufig (*nie bis ein mal pro Woche*) verschiedene handlungs- und produktionsorientierte Herangehensweisen an erzählende und dramatische Texte im Deutschunterricht realisiert werden. Als Beispielitem sei hier ein typisches Item der Skala vorgestellt: „Kommt es vor, dass Sie im Deutschunterricht die folgenden Verfahren einsetzen? Bei erzählenden und dramatischen Texten: Figuren in Ich-Form vorstellen“. 20% der Deutschlehrkräfte gaben an, dieses Verfahren ein paar mal im Monat in ihrem Unterricht einzusetzen, 60% ein paar mal im Jahr und 20% gaben an, nie darauf zurückgreifen. Weitere Operationalisierungen didaktischer Unterrichtselemente zu Prosa erfolgten über die Frage nach szenischer Umsetzung von gelesenen Texten, der Darstellung von Personen aus der Ich-Perspektive und der Umsetzung in andere Textsorten. Insgesamt zeigt sich, dass die didaktisch kreative Herangehensweisen an Prosatexte für einen Großteil der Deutschlehrkräfte nur zum gelegentlich eingesetzten Repertoire ihres Unterrichts gehört.

Unterrichtselemente Poesie

Die Skala „Unterrichtselemente Poesie“ ($\alpha = .91$) erfasst mit ebenfalls 15 Items die Häufigkeit, mit der sich Deutschlehrkräfte aus einem vorgegebenen Repertoire er-

fragter handlungs- und produktionsorientierter Elemente bei der Annäherung an Lyrik bedienen. Ein typisches Item dieser Skala lautet: „Kommt es vor, dass Sie im Deutschunterricht die folgenden Verfahren einsetzen? Bei Gedichten: weggelassene Wörter oder Zeilen ergänzen.“ 16% der Deutschlehrkräfte greifen mindestens ein paar mal pro Monat auf diese Form der Arbeit mit Gedichten zurück; der weitaus größte Teil der Lehrerinnen und Lehrer (62%) jedoch nur ein paar mal im Jahr und 22% sogar nie.

Dieses Ergebnis ist analog zur Erarbeitung von Prosatexten im Deutschunterricht: Auch andere didaktische Zugänge zu Gedichten, wie die Entflechtung von Gedichten, das eigene Verfassen eines Gedichts und das Umschreiben in eine andere Textsorte zählen für die meisten Deutschlehrkräfte nur in größeren zeitlichen Abständen zu ihrem Unterrichtsalltag.

Schreibförderliche Unterrichtselemente

Die letzte im Zusammenhang mit literarischem Unterricht in DESI eingesetzte Skala, die hier vorgestellt werden soll, erfasst den Einsatz von Unterrichtselementen, die die Schreibfähigkeit der Schülerinnen und Schüler fördern sollen. Die Skala „schreibförderliche Unterrichtselemente“ ($\alpha = .78$), die aus 16 Items besteht, erfasst nicht die Bedeutung sprachbezogener Kompetenzen im Deutschunterricht (vgl. dazu den folgenden Abschnitt), sondern die Förderung des kreativen Potentials der Schülerinnen und Schüler. Ein typisches Item der Skala „schreibförderliche Unterrichtselemente“ lautet: „Wie oft setzen Sie die folgenden Unterrichtsmethoden zur Förderung des Kompetenzbereichs Schreiben in Ihrer Klasse/ Ihrem Kurs ein? Schreibkonferenzen/ Schreibwerkstatt durchführen“. Im Einklang mit den Ergebnissen der Gesamtskala gibt die Hälfte der Deutschlehrkräfte (52%) an, nie eine Schreibwerkstatt mit ihren Schülerinnen und Schülern zu veranstalten. 42% machen dies ein paar mal pro Jahr. Für den größten Teil der Lehrkräfte ist die systematische Förderung der kreativen Schreibfähigkeiten nicht Teil der regulären, alltäglichen Praxis ihres Unterrichts.

Bedeutung sprachbezogener Kompetenzen

Welche Bedeutung hat die Förderung sprachbezogener Kompetenzen im Deutschunterricht? Um dieser Frage nachzugehen, wurde in DESI die Implementierung bestimmter, auf die Förderung sprachlicher Kompetenzen bezogener Unterrichtselemente erfragt. Im Einzelnen gaben die Deutschlehrkräfte auf einer vierstufigen Antwortskala (*nie, ein paar mal pro Jahr, ein paar mal pro Monat, ein paar mal pro Woche*) Auskunft in den Skalen „Wortschatz im Unterricht“, „Sprachbewusstheit im Unterricht“, „Grammatikförderliche Unterrichtselemente“, „Rechtschreibförderliche Unterrichtselemente“, „Lernsituationen Leseverstehen“ und „Lernsituationen Rechtschreibung“. Die Iteminhalte dieser Skalen sind in Tabelle 28.4 dargestellt. Wenn die Lehrkräfte angaben, sprachbezogene Kompetenzen *ein paar Mal pro Woche* zu fördern, wurde dies als *routinemäßig* angewendetes Element des Deutschunterrichts angesehen.

Tabelle 28.4: Förderung sprachbezogener Kompetenzen im Deutschunterricht.

Skalen- und Iteminhalte	% routine- mäßig
Skala: Wortschatz im Unterricht (4 Items; $\alpha = .80$)	
Wie häufig wurden in diesem Schuljahr (9. Klasse) in Ihrem Unterricht die folgenden Aufgabentypen behandelt?	
- über die Bedeutung verschiedener Wörter sprechen	15
- über verschiedene Möglichkeiten sprechen, etwas Ähnliches mit unterschiedlichen Wörtern auszudrücken (...)	12
- darüber sprechen, wie man etwas noch präziser und passender ausdrücken kann	24
- die Bezeichnung verschiedener Gegenstände üben (...)	4
Skala: Sprachbewusstheit im Unterricht (5 Items; $\alpha = .84$)	
Wie häufig wurden in diesem Schuljahr (9. Klasse) in Ihrem Unterricht die folgenden Aufgabentypen behandelt?	
- wie sich gesprochene und geschriebene Sprache voneinander unterscheiden	9
- dass verschiedene Personengruppen (...) unterschiedlich sprechen	6
- dass man bei verschiedenen Anlässen (...) unterschiedlich schreibt	6
- wie sich direkte (...) und indirekte Rede (...) unterscheiden	4
- wie man den Konjunktiv verwendet (Wenn er es gewusst hätte, wäre er hier geblieben)	4
Skala: Grammatikförderliche Unterrichtselemente (3 Items, $\alpha = .86$)	
Wie oft setzen Sie die folgenden Unterrichtsmethoden zur Förderung des Kompetenzbereichs <u>Sprache und Sprachgebrauch untersuchen</u> in Ihrer Klasse / Ihrem Kurs ein?	
- Grammatische Proben anwenden (...)	8
- Rechtschreibstrategien anwenden (...)	16
- Nachschlagewerke nutzen	17
Skala: Rechtschreibförderliche Unterrichtselemente (9 Items, $\alpha = .86$)	
Wie oft tun Sie die folgenden Dinge?	
- der Klasse/ dem Kurs Rechtschreibregeln erläutern	21
- die Schüler/innen mit Lernkarteien üben lassen	1
- Fehler aus Aufsätzen oder freien Texten der Schüler/innen besprechen	13
- den Schülern/innen Zeit geben, individuell an ihren Fehlerschwerpunkten zu arbeiten	8
- den Schülern/innen verschiedene Lerntechniken erklären oder vormachen	8
- die Schüler/innen Übungsdiktate schreiben lassen	6
- Übungen zur Rechtschreibung als Hausaufgaben aufgeben	9
- den Schülern/innen Gelegenheit zur Selbstkontrolle des Geschriebenen geben	9
- die Schüler/innen in Partnerarbeit üben lassen	7
Skala: Lernsituationen Leseverstehen (7 Items; $\alpha = .83$)	
Wie häufig wurden in diesem Schuljahr (9. Klasse) in Ihrem Unterricht die folgenden Aufgabentypen behandelt? Einen Text lesen und anschließend...	
- erraten, woher dieser Text stammen könnte (z.B. aus einer Zeitung, einem Roman usw.)	1
- verschiedene Fragen zum Inhalt beantworten	38
- bestimmte Textstellen ausschreiben	19
- verschiedene Textstellen miteinander vergleichen (...)	14
- verschiedene Meinungen zu dem Text miteinander vergleichen und beurteilen	15
- bestimmte Textstellen unterstreichen	35
- Schlüsselwörter ausschreiben bzw. am Textrand notieren	27
Skala: Lernsituationen Rechtschreibung (3 Items, $\alpha = .85$)	
Wie häufig wurden in diesem Schuljahr (9. Klasse) in Ihrem Unterricht die folgenden Aufgabentypen behandelt? Rechtschreibfehler besprechen, und zwar ...	
- bei der Besprechung der Hausaufgaben	23
- beim Schreiben an der Tafel	45
- wenn im Unterricht schwierige Worte auftauchen	45

Kooperation

Kooperation von Lehrkräften wird traditionell als wichtiger Aspekt von Schulqualität angesehen; die empirischen Befunde zum Zusammenhang von Lehrerkooperation und Schülerleistung sind heterogen (vgl. Kap. 34). Die Deutschlehrkräfte wurden in DESI gefragt, wie häufig sie ihren Unterricht gemeinsam vorbereiten und gemeinsam durchführen. Beide Formen der unterrichtsbezogenen Kooperation sind in Integrierten Gesamtschulen am häufigsten verbreitet; immerhin 21% der Klassen werden hier von Lehrkräften unterrichtet, die ihren Unterricht mehrmals pro Monat gemeinsam planen. Danach folgen mit abnehmender Häufigkeit die Hauptschulen, die Realschulen und schließlich der gymnasialen Bildungsgang, dessen Lehrkräfte vergleichsweise wenig kooperieren. Über alle Schulformen hinweg bereiten die Lehrkräfte von 8% aller Klassen ihren Unterricht gemeinsam vor, aber nur 2% insgesamt führen den Deutschunterricht gelegentlich gemeinsam durch.

Integrierter Blick auf Grunddimensionen der Alltagspraxis des Deutschunterrichts

Nachdem verschiedene Aspekte des Unterrichts einzeln dargestellt wurden, stellt sich die Frage, ob sich die Perspektive der Lehrkräfte durch gemeinsame, übergeordnete Basisdimensionen beschreiben lässt, ob also geteilte Vorstellungen von grundlegenden Aspekten der Alltagspraxis des Deutschunterrichts bestehen. Mittels konfirmatorischer Faktorenanalyse lässt sich ein dreidimensionales Modell anpassen, in dem sich die folgenden drei latenten Dimensionen abbilden lassen³:

- die Vielfalt der Unterrichtsmethoden
- literarischer Unterricht
- sprachbezogener Unterricht

Inwieweit dem Ideal eines abwechslungsreich gestalteten Unterrichts in der Alltagspraxis des Deutschunterrichts entsprochen wird, lässt sich anhand der Dimension „Vielfalt der Unterrichtsmethoden“ abbilden: Deutschlehrkräfte unterscheiden sich im Ausmaß, in dem sie moderne Unterrichtsmethoden anwenden. Die zweite Dimension „literarischer Unterricht“ differenziert zwischen den Lehrkräften hinsichtlich der Bedeutung, die sie einem produktiven literarischen Unterricht in ihrer eigenen Alltagspraxis zumessen. Hier fließt sowohl der Stellenwert von literarischen Unterrichtsmaterialien als auch die Ausrichtung des Unterrichts auf literarische Inhalte sowie die Förderung des eigenen kreativen Schreibens ein. Die dritte Dimension der Alltagspraxis des Deutschunterrichts bezieht sich auf die Bedeutung sprachbezogener Komponenten im Unterricht. Lehrkräfte, die auf dieser Dimension eine hohe Ausprägung haben, legen, einem utilitaristischen Paradigma des Muttersprachunterrichts entsprechend, besonderen Wert darauf, ihre Schülerinnen und Schüler für die alltägliche Kommunikation auszurüsten. Dabei stehen mündliche und schriftliche Kommunikation im Mittelpunkt des Deutschunterrichts. Die

3 CFI = .875; RMSEA = 0.073

drei genannten Dimensionen bilden ein differenziertes Bild von Unterricht ab, auch wenn sie untereinander korrelieren. Die allgemeindidaktische Dimension „Vielfalt der Unterrichtsmethoden“ korreliert (auf Klassenebene) vergleichsweise gering mit beiden anderen Dimensionen ($r = .24$ mit Dimension 2; $r = .31$ mit Dimension 3), während die beiden eher fachspezifische Aspekte von Unterrichtsqualität abbildenden Dimensionen „literarischer Unterricht“ und „sprachbezogener Unterricht“ höher miteinander korrelieren ($r = .41$). Tabelle 28.5 gibt einen Überblick über die drei Dimensionen von Unterrichtsqualität aus Sicht der Lehrkräfte, die zugehörigen in vorhergehenden Abschnitten dargestellten Skalen und die entsprechenden Faktorladungen.

Tabelle 28.5. Faktorladungen des dreidimensionalen Modells von Unterrichtsqualität aus Sicht von Lehrerinnen und Lehrern. Dargestellt sind die Skalen mit ihren jeweiligen Ladungen auf dem latenten Konstrukt, dem sie zugeordnet sind⁴.

Faktoren und Indikatoren (Skalen des Lehrerfragebogens)	Ladungen
Dimension 1: Vielfalt der Unterrichtsmethoden	
Partizipation	0.71
Differenzierung	0.69
Alternative Unterrichtsmethoden	0.68
Kleingruppenarbeit	0.61
Dimension 2: literarischer Unterricht	
Unterrichtselemente Poesie	0.88
Unterrichtselemente Prosa	0.85
Schreibförderliche Unterrichtselemente	0.57
Unterrichtsmaterialien Literarische Texte	0.48
Dimension 3: sprachbezogener Unterricht	
Wortschatz im Unterricht	0.71
Grammatikförderliche Unterrichtselemente	0.70
Lernsituationen Rechtschreibung	0.65
Rechtschreibförderliche Unterrichtselemente	0.61
Sprachbewusstheit im Unterricht	0.58
Lernsituationen Leseverstehen	0.58

⁴ Alle Nebenladungen sind $< .25$.

Kann man Deutschlehrkräfte nun danach unterscheiden, welche relative Bedeutung sie diesen drei Dimensionen beimessen? In Analogie zu anderen Studien (PISA, vgl. Klieme/Rakoczy, 2003; Baumert u.a. 2004,) ist es naheliegend zu erwarten, dass es verschiedene Befunde je nach Schulform gibt. Um ein vollständigeres Bild des Unterrichtsalltags zu erhalten, werden an dieser Stelle die Häufigkeit von Lernerfolgskontrollen und der didaktischen Stellenwert von Hausaufgaben sowie die Kooperation der Lehrkräfte bei der Unterrichtsvorbereitung und -durchführung mit herangezogen. Abbildung 28.3 veranschaulicht die Schulformprofile grafisch.

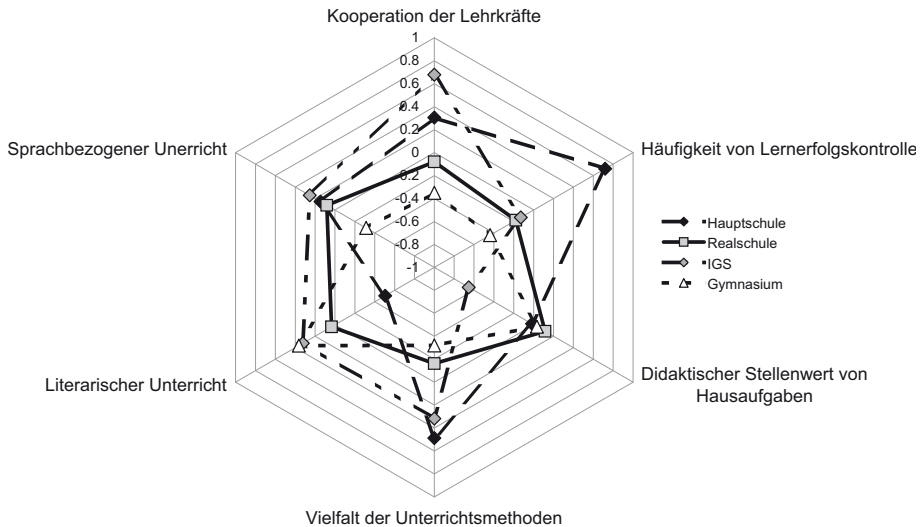


Abbildung 28.3. Ausprägungen grundlegender Merkmale des Deutschunterrichts aus Sicht von Lehrkräften nach Bildungsgängen, z-standardisierte Skalenmittelwerte.

Die Unterrichtspraxis der verschiedenen Bildungsgänge aus Sicht der Lehrkräfte weist tatsächlich unterschiedliche Profile auf. Hinsichtlich der inhaltlichen Schwerpunkte zeigen Hauptschule (Fokus: Sprache) und Gymnasium (Fokus: Literatur) entgegengesetzte Profile. Der Deutschunterricht in der Hauptschule zeichnet sich darüber hinaus im Vergleich zu anderen Bildungsgängen durch eine größere Vielfalt der angewandten Unterrichtsmethoden und eine dichtere Abfolge von Lernerfolgskontrollen aus. Spezifisches Merkmal der Integrierten Gesamtschulen ist nach Angaben der Lehrkräfte eine ausgeprägte Kooperation, die geteilte Unterrichtsvorbereitung ebenso einschließt wie (hin und wieder) gemeinsames Unterrichten. Hausaufgaben spielen hier im Unterrichtsgeschehen allerdings eine untergeordnete Rolle. Auffallend ist schließlich, dass im gymnasialen Bildungsgang weder auf Methodenvielfalt im Unterricht noch auf Teamarbeit im Kollegium ein starker Wert gelegt wird.

28.2 Deutschunterricht aus Sicht von Schülerinnen und Schülern

Aus der Unterrichtsforschung ist bekannt, dass Jugendliche zentrale Qualitätsdimensionen von Unterricht valide einschätzen können, vor allem Aspekte der Klassenführung und Strukturiertheit sowie das Unterrichtsklima (vgl. etwa Clausen 2002). Angesichts des Fehlens von Videoaufzeichnungen müssen sich Analysen zur Wirksamkeit des Deutschunterrichts wesentlich auf die Urteile der Schülerinnen und Schüler stützen, die sehr differenziert erfasst wurden. Drei Typen von Fragebogenskalen wurden eingesetzt (vgl. Tabelle 28.6).

Insgesamt elf Skalen erfassten *Aspekte der Unterrichtsqualität*, die in der Unterrichtsforschung seit langem erforscht und als relevant nachgewiesen wurden, von der Klassenführung über Motivierung und Fehlerkultur bis zur Verständlichkeit des Unterrichts. Den Schülern wurden jeweils zwischen 2 und 14 Aussagen vorgegeben, die auf einer vierstufigen Skala (von *stimmt gar nicht* bis *stimmt ganz genau*) einzuschätzen waren. Die Aussagen beziehen sich explizit auf den Deutschunterricht, betreffen aber allgemeine Unterrichtsmerkmale, und sind bis auf den Bezug zum Fach Deutsch identisch mit den für die Englischunterricht entworfenen Skalen (siehe Kap. 32).

Mit drei Skalen wurde erfragt, welche *Zielsetzungen* die Schülerinnen und Schüler bei ihren Lehrpersonen wahrnehmen. Dabei wurde versucht, zu unterscheiden, inwieweit die Lehrperson mündliche Fähigkeiten, sprachsystematische Kompetenzen sowie Lerntechniken fördern. Hier war unter vier Antwortalternativen (von *unwichtig* bis *sehr wichtig*) zu wählen.

Mit sechs Skalen zu jeweils 2 bis 10 Items wurde schließlich erfragt, wie häufig im Unterricht *spezifische Lerngelegenheiten* vorkamen, die den Anforderungen der DESI-Subtests für das Fach Deutsch entsprechen. Jede einzelne Aussage war auf einer vierstufigen Skala (von *nie* bis *ein paar mal pro Woche*) einzustufen. Die Auflistung der relevanten Lernsituationen war mit den Fachdidaktikern abgestimmt, welche die jeweiligen Tests entwickelt hatten.

Tabelle 28.6: Skalen zur Qualität des Deutschunterrichts aus Schülersicht⁵.

Skala	Itemzahl	α	Beispielitem
Aspekte der Unterrichtsqualität			
Unterstützung	3	.83	Wenn ich Hilfe brauche, bekomme ich sie von meinem Deutschlehrer.
Schülerorientierung	6	.87	Wenn jemand eine gute Idee hat, dann geht mein Deutschlehrer darauf ein.
Motivierung, thematisch	2	.79	Mein Deutschlehrer kann mich manchmal richtig für die Unterrichtsthemen begeistern.
Verständlichkeit	3	.76	Die Aufgabenstellungen im Deutschunterricht sind für mich klar und verständlich.
Unterrichtsklima	3	.83	Ich komme mit meinem Deutschlehrer gut aus.
Positive Fehlerkultur	14	.89	Aus der Verbesserung meiner Fehler durch meinen Deutschlehrer lerne ich etwas dazu.
Strukturiertheit des Unterrichts	3	.77	Mein Deutschlehrer gibt Hinweise, worauf es in der Unterrichtsstunde besonders ankommt.
Motivierung, instrumentell	3	.83	Mein Deutschlehrer betont, dass gutes Deutsch in vielen Berufen eine große Rolle spielt.
Klassenführung	2	.77	Mein Deutschlehrer sorgt dafür, dass die Schüler die ganze Stunde über aufpassen.
Strukturierung Gruppenarbeit	3	.78	Nach einer Gruppenarbeit werden im Deutschunterricht die Ergebnisse der Gruppen vorgetragen.
Tempodruck	4	.75	Der Deutschstoff wird für mich zu schnell durchgenommen.
Zielsetzungen der Lehrkraft			
Wie wichtig ist deinem Deutschlehrer Folgendes?			
Wichtigkeit sprachbezogener Kompetenzen	10	.89	- dass ihr grammatikalisch richtig sprechen könnt.
Wichtigkeit von Lerntechniken	7	.80	- dass ihr Lerntechniken (wie Markieren von Schlüsselstellen in Texten) einsetzen könnt.
Wichtigkeit mündlicher Fähigkeiten	3	.67	- dass eure Aussprache deutlich und korrekt ist.
Spezifische Lerngelegenheiten			
Wie häufig wurden in diesem Schuljahr die folgenden Dinge behandelt?			
Lernsituationen Wortschatz	4	.80	- darüber sprechen, wie man etwas noch präziser und passender ausdrücken kann.
Lernsituationen Sprachbewusstheit	5	.78	- darüber sprechen wie sich gesprochene und geschriebene Sprache voneinander unterscheiden.
Lernsituationen Briefe schreiben	2	.72	- einen privaten Brief schreiben.
Lernsituationen Leseverstehen	7	.83	- bestimmte Textstellen unterstreichen
Lernsituationen Kommunikation und Argumentation	10	.89	- Formulieren von Gegenargumenten zu verschiedenen Argumenten, z.B. zu einem Pro-Argument ein Contra-Argument
Lernsituationen Rechtschreibung	3	.81	- Rechtschreibfehler besprechen, und zwar beim Schreiben an der Tafel?

Alle untersuchten Skalen besitzen, wie Tabelle 28.6 ausweist, eine ausreichende, häufig gute bis sehr gute Konsistenz: Der Median der α -Koeffizienten liegt bei .80.

⁵ Hier ist aus Platzgründen nur von Lehrern die Rede; im Schülerfragebogen wird dagegen durchgehend geschlechtsneutral formuliert (dein Lehrer/ deine Lehrerin).

Ähnlich wie bei den Urteilen der Lehrkräfte gilt es zu untersuchen, ob sich die 20 Skalen auf einige wenige Grunddimensionen zurückführen lassen.

Explorative Analysen zeigten, dass die drei oben unterschiedenen Typen von Skalen getrennte Grunddimensionen ergeben, wobei sich die Skalen zur allgemeinen Unterrichtsqualität sogar noch weiter unterteilen lassen. Entgegen der Hoffnung, in Anlehnung an die differenzierte DESI-Testbatterie auch unterschiedliche Unterrichtsprofile identifizieren zu können, ergaben sich jedoch keine Teil-Faktoren im Gefüge der erfassten Lerngelegenheiten. Vielmehr war festzustellen, dass Schülerinnen und Schüler, die beispielsweise davon berichteten, im Unterricht werde häufig das Schreiben von Briefen verlangt oder geübt, auch angaben, es würden häufig Rechtschreibfehler besprochen. Alle Subtest-spezifischen Skalen für einschlägige Lernsituationen ergeben somit einen einzigen Faktor, auf dem ein Unterricht hohe Werte erzielt, wenn quer über die Anforderungsbereiche hinweg eine Vielzahl von Lernsituationen häufig realisiert werden. Dieser Faktor lässt sich daher als „Vielfalt von Lerngelegenheiten für den Erwerb sprachlicher Kompetenzen“ charakterisieren.

In ähnlicher Weise ergibt sich ein Faktor, der alle drei Skalen der wahrgenommenen Zielsetzungen von Deutschlehrkräften umfasst. Auch hier lassen sich also aus den Wahrnehmungen der Schülerinnen und Schüler keine differenzierten Profile ableiten, sondern es wird ein einziger Faktor gebildet, der etwas darüber aussagt, inwieweit die Jugendlichen grundsätzlich wahrnehmen, dass die Lehrperson zielorientiert vorgeht und klare Leistungsansprüche stellt. Dieser Faktor lässt sich am ehesten als „wahrgenommenes Anspruchsniveau im Bereich sprachbezogener Kompetenzen“ interpretieren. Hier wie auch bei den Lernsituationen blieben Zielbereiche des Deutschunterrichts, die nicht in den DESI-Tests abgebildet waren, wie etwa literatur- und kulturgeschichtliche Aspekte, unberücksichtigt.

Die allgemeinen Unterrichtsmerkmale lassen sich recht gut in zwei Dimensionen gliedern⁶: Die eine deckt mit „Unterrichtsklima“, „Schülerorientierung“, „Unterstützung“ und verwandten Merkmalen die Grunddimension der Unterrichtsqualität ab, die bei Klieme u.a. (2006) als „schülerorientiertes Unterrichtsklima“ bezeichnet wird. Die andere umfasst mit den Skalen Klassenführung, Strukturiertheit des Unterrichts und Strukturiertheit der Gruppenarbeit Aspekte, die in der Unterrichtsforschung mit der Grunddimension der „effizienten und strukturierten Klassenführung“ verknüpft werden. Strukturiertheit stellt, wie Rakoczy u.a. (2007) zeigen, ein Unterrichtsmerkmal dar, dass je nach Operationalisierungsansatz und Breite der vorfindlichen Praxis teils der effizienten Klassenführung verwandt ist, teils der kognitiven Aktivierung. Im hier untersuchten Deutschunterricht muss offensichtlich *Strukturierung* als *Aspekt einer effizienten Klassenführung* verstanden werden. Außerdem gehört zu dieser Grunddimension hier die „instrumentelle“, auf außerschulische Ziele gerichtete Motivierung, die in der Tat als Merkmal eines klar zielorientierten Unterrichts verstanden werden kann.

Bezogen auf das Grundmodell von Unterrichtsqualität, das Klieme u.a. (2006) beschreiben, bleibt zu klären, wo in diesem Modell der Unterrichtsqualität für das

6 Die Variable Tempodruck lässt sich keinem übergeordneten Faktor zuordnen.

Fach Deutsch so etwas wie „kognitive Aktivierung“ steckt. Diese Grunddimension ist, da sie im Wesentlichen den fachdidaktischen Anforderungscharakter repräsentiert, in Schülerurteilen nur schwer zu erkennen. Mit einiger Vorsicht interpretieren wir die beiden hier gefundenen Dimensionen „Vielfalt der Lerngelegenheiten“ sowie „Anspruchsniveau“ als Teilaspekte einer im weitesten Sinne kognitiv aktivierenden Unterrichtsführung.

Die beschriebene Grundstruktur mit vier Dimensionen ließ sich in einer konfirmatorischen Faktorenanalyse mit sehr guten Anpassungswerten empirisch nachweisen⁷. Dabei wurde im Sinne eines Zwei-Ebenen-Modells zwischen der Struktur der innerhalb einer Klasse geteilten Wahrnehmungen (Klassenebene) einerseits und der Struktur der individuellen Wahrnehmungstendenzen (individuelle Ebene) andererseits unterschieden, wobei die Analyse jedoch davon ausging, dass beide Strukturen parallel aufgebaut sind. Tatsächlich lässt sich zeigen, dass die vier beschriebenen Grunddimensionen der Unterrichtsqualität im Fach Deutsch sowohl die innerhalb einer Schulklasse geteilte Wahrnehmung als auch die individuellen Wahrnehmungsmuster beschreiben können.

Tabelle 28.7 dokumentiert, dass alle 20 Teilskalen ausreichend viel Varianz auf der Klassenebene binden (gemessen an der Intraklassenkorrelation (ICC)), so dass sie in der Tat als differenzierende Charakteristika der Lernumgebung behandelt werden können. Dies gilt besonders für die allgemeinen Unterrichtsmerkmale (Unterrichtsklima bzw. Strukturiertheit), während es hinsichtlich des wahrgenommenen Anspruchsniveaus und der wahrgenommenen Lerngelegenheiten offensichtlich eine größere Variation innerhalb der Schulklassen gibt.

7 $RMSEA = 0.04$; $SRMR_{between} = 0.10$; $SRMR_{within} = 0.04$; $CFI = 0.95$

Tabelle 28.7: 4-Faktorenmodell der Unterrichtsqualität aus Sicht von Schülerinnen und Schülern mit Intraklassenkorrelationen (ICC), Faktorladungen (λ) und Varianzen (Var.) der latenten Dimensionen.

Skalen/Indikatoren	ICC	λ Indivi- dualebene	λ Klassenebene
Dimension 1: Schülerorientiertes Unterrichtsklima		Var = .28	Var = .08
Unterstützung	.19	.83	.97
Schülerorientierung	.20	.83	.95
Thematische Motivierung	.19	.67	.92
Verständlichkeit	.18	.77	.92
Unterrichtsklima	.16	.70	.93
Fehlerkultur	.16	.79	.96
Dimension 2: Strukturierte Klassenführung		Var = .14	Var = .06
Strukturiertheit des Unterrichts	.16	.57	.79
Instrumentelle Motivierung	.13	.63	.82
Klassenführung	.22	.60	.81
Strukturierung Gruppenarbeit	.18	.62	.70
Dimension 3: Anspruchsniveau		Var = .25	Var = .06
Wichtigkeit sprachbezogener Kompetenzen	.15	.85	.58
Wichtigkeit von Lerntechniken	.14	.71	.84
Wichtigkeit mündlicher Fähigkeiten	.14	.66	.71
Dimension 4: Vielfalt von Lerngelegenheiten		Var = .20	Var = .01
Lernsituationen Wortschatz	.12	.74	.98
Lernsituationen Sprachbewusstheit	.13	.78	.87
Lernsituationen Briefe schreiben	.14	.47	.64
Lernsituationen Leseverstehen	.11	.78	.64
Lernsituationen Kommunikation	.12	.83	.74
Lernsituationen Rechtschreibung	.20	.65	.93

Auch auf der Basis der Urteile von Schülerinnen und Schülern soll abschließend gefragt werden, inwieweit sich in den unterschiedlichen Bildungsgängen *differenzielle Entwicklungsmilieus* (Baumert u.a. 2004) zeigen. Hierzu sind die mittleren Einschätzungen der vier latenten Dimensionen von Unterrichtsqualität sowie zusätzlich das Merkmal Tempodruck in Abbildung 28.4 grafisch dargestellt. Am deutlichsten fällt auf, dass – vergleichbar dem Faktor „Vielfalt der Unterrichtsmethoden“ im Lehrerurteil – Hauptschule und Gymnasium ein entgegengesetztes Profil in Bezug auf die „Vielfalt der Lerngelegenheiten“ besitzen: Auch gemessen an der Wahrnehmung der Schülerinnen und Schüler bietet der Gymnasialunterricht, in der Tendenz auch der Realschulunterricht, weniger vielfältige Lernsituationen als der Unterricht an Gesamtschulen und vor allem an Hauptschulen. Der wahrgenommene Tempodruck ist hingegen in Gesamtschule und Hauptschule vergleichsweise hoch, in der Realschule und vor allem im Gymnasium niedrig. In der Tendenz zeichnen sich Haupt- und Gesamtschule auch durch etwas größere Strukturiertheit und ein etwas stärker schülerorientiertes Unterrichtsklima aus. Relativ geringe Unterschiede

gibt es interessanterweise im wahrgenommenen Anspruchsniveau –erklärbar mit der Tatsache, dass Schüler aller Schulformen sich hierbei an ihnen jeweils vertrauten Lernkontexten orientieren und als Referenz nicht den gesamten Alterjahrgang in allen Bildungsgängen zugrunde legen.

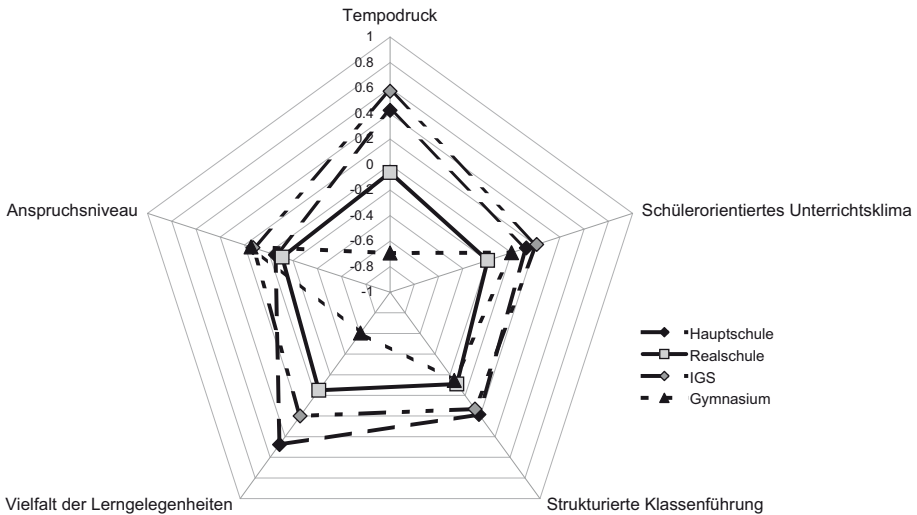


Abbildung 28.4: Ausprägungen der geteilten Unterrichtswahrnehmung von Schülerinnen und Schülern nach Bildungsgängen, z-standardisierte Skalenmittelwerte.

Der Befund, wonach gerade der gymnasiale Deutschunterricht aus Schülersicht eine offensichtlich weniger anregungsreiche Lernumwelt darstellt als der Deutschunterricht anderer Schulformen, entspricht den Ergebnissen, die Baumert u.a. (2004) aus PISA 2003 für den Mathematikunterricht berichten. Dies rechtfertigt für nachfolgende Forschung die Hypothese, dass der Unterricht im Gymnasium, auch im Fach Deutsch, didaktisch weniger anregungsreich gestaltet ist, somit das Potenzial der Schülerinnen und Schüler möglicherweise nicht voll ausschöpft. Damit ist das Verhältnis von Unterrichtsmerkmalen, Lernvoraussetzungen und Lernergebnissen angesprochen, das im Folgenden empirisch untersucht werden soll.

28.3 Adaptivität und Wirksamkeit des Deutschunterrichts: Zusammenhänge zwischen wahrgenommener Unterrichtsqualität und Kompetenz

Das Design der DESI-Studie ermöglicht es, die Veränderung der Kompetenz im Deutschen, d. h. den Leistungszuwachs während der 9. Jahrgangsstufe, in den Bereichen Lesen und Sprachbewusstheit abzuschätzen. Im Folgenden werden die Zusammenhänge zwischen den verschiedenen Aspekten der wahrgenommenen Unterrichtsqualität und der erfassten Leistungsveränderung analysiert; wir fol-

gen dabei dem in Kapitel 26 beschriebenen Vorgehen. In Mehr-Ebenen-Modellen, die jeweils Effekte auf Schülerebene und auf Klassenebene simultan in die Auswertung einbeziehen, wurden alle Skalen zur Unterrichtsqualität jeweils einzeln in Regressionsanalysen aufgenommen, wobei der Kompetenzstand am Ende der 9. Jahrgangsstufe aus dem Kompetenzstand am Anfang sowie dem betreffenden Unterrichtsqualitätsmerkmal sowie weiteren Faktoren erklärt wurde. Als Hintergrundmerkmale wurden dabei durchweg der Bildungsgang, der individuelle Sprachhintergrund (Erstsprache) und der Anteil der Schüler mit nicht deutscher Erstsprache, die kognitiven Grundfähigkeiten des einzelnen Schülers und der einzelnen Schülerin sowie der entsprechende Klassendurchschnitt, das Geschlecht sowie der Mädchenanteil in der Klasse, der sozioökonomische Status der Herkunftsfamilie und die sozioökonomische Komposition der Klasse einbezogen.

Die Mehrebenen-Pfadanalysen, die schematisch in Abbildung 28.5 dargestellt sind, erlauben es, systematisch zwei Fragestellungen zu trennen:

- Inwieweit ist Deutschunterricht adaptiv? Untersucht wird, ob unter Berücksichtigung der genannten Rahmenbedingungen der Deutschunterrichts dem Leistungsniveau ihrer Klassen zu Beginn der 9. Jahrgangsstufe angepasst ist (Pfad A in Abb. 28.5), beispielsweise durch die Schaffung spezifischer Lernsituationen, durch ein entsprechendes Anspruchsniveau oder durch Unterrichtsklima und Klassenführung.
- Welche Auswirkungen hat die Unterrichtsgestaltung auf den Kompetenzzuwachs unter Kontrolle von Lernvoraussetzungen und diversen Hintergrundmerkmalen (Pfad B in Abb. 28.5), hier wird geprüft, ob das jeweilige Unterrichtsmerkmal (in Abb. 28.5 mittlerer Kasten) über die Bedeutung der jeweiligen Lernvoraussetzungen und der Rahmenbedingungen hinaus einen Beitrag zur Erklärung der Lernergebnisse leisten kann.

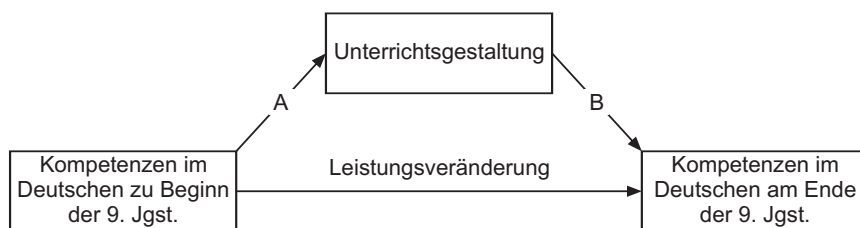


Abbildung 28.5 Grafische Darstellung der Mehrebenen-Regressionsanalysen zum Einfluss von Unterrichtsqualität auf die Veränderung der Deutschkompetenz.

Verschiedene Merkmale der Unterrichtsgestaltung – teils aus Lehrer-, teils aus Schülersicht – wurden nach diesem Muster unabhängig voneinander ausgewertet. Tabelle 28.8 gibt einen Überblick über die Zusammenhänge zwischen den verschiedenen Aspekten der wahrgenommenen Unterrichtsqualität und der erfassten Leistung (Lernvoraussetzung bzw. Lernresultat).

Tabelle 28.8 Zusammenhänge zwischen den verschiedenen Aspekten der wahrgenommenen Unterrichtsqualität und der erfassten Leistung. Standardisierte Regressionskoeffizienten auf Klassenebene unter Kontrolle der im Text genannten Hintergrundvariablen.

	Adaptivität des Unterrichts: Unterrichtsmerkmal in Abhängigkeit von der Eingangsleistung (Pfad A)		Wirkung des Unterrichts: Kompetenzentwicklung in Abhängigkeit vom Unterrichtsmerkmal (Pfad B)	
	Lese- kompetenz	Sprach- bewusstheit	Lese- kompetenz	Sprach- bewusstheit
Allgemeine Unterrichtsmerkmale (Schülersicht)				
Tempodruck	-.52*	-.50*	-.18*	-.11*
Anspruchsniveau	.32*	.32*	.11*	.09*
Disziplin im Unterricht	.23**	.21**	.04	.04*
Klassenführung	.00	-.07	.03	.03
Strukturiertheit	-.28*	-.34*	.01	-.02
Verständlichkeit	.06	.07	.07*	.03
Schülerorientierung	.00	.00	.01	.04*
Unterstützung	-.11*	-.13*	.03	.03
Unterrichtsklima	.00	-.07	.02	.02
Gruppenarbeit	.11*	.07	.03	.04
Hausaufgaben	.00	.00	.06	.05*
Spezifische Lerngelegenheiten (Schülersicht)				
Wortschatz	-.33*	-.33*	.02	.03
Sprachbewusstheit	-.33*	-.33*	.04	.03
Briefe schreiben	-.33*	-.40*	-.01	.01
Rechtschreibung	-.38*	-.46*	.04	.04
Fachdidaktische Merkmale (Lehrersicht)				
Alternative Unterrichtsmethoden	-.33*	-.32*	-.05	.02
Rechtschreibförderliche Unterrichtselemente	-.32*	-.40*	.02	.02
Grammatikförderliche Unterrichtselemente	-.27*	-.27*	.02	-.05
Unterrichtselemente Prosa	.26*	.33*	.07*	.00

Wie die erste Zeile der Tab. 28.8 belegt, besteht – nach Anrechnung aller Hintergrundvariablen auf Individual- und Klassenebene – ein besonders enger Zusammenhang zwischen dem Leistungsniveau der Klasse – egal, ob es durch den Lesetest oder den Test zur Sprachbewusstheit operationalisiert ist – und dem wahrgenommenen Tempodruck. Je niedriger das Leistungsniveau zu Beginn des Schuljahres, umso stärker wird der Tempodruck erlebt, und je größer das wahrgenommene Unterrichtstempo, desto geringer ist der Lernzuwachs bis zum Ende des Schuljahres. Diese Aussagen verweisen weniger auf einer echten „Wirkfaktor“ im Deutschunterricht als dass sie die Validität des Schülerurteils belegen: Weist eine Klasse gemessen an ihren Rahmenbedingungen auffällig schwache Leistungen auf, so spiegelt sich dies in dem Gefühl der Überforderung bzw. der Wahrnehmung eines nicht mehr zu bewältigenden Unterrichtstempos, und umgekehrt ist ein als besonders hoch erlebtes

Unterrichtstempo ein Hinweis auf den im Schuljahresverlauf vergleichsweise geringen Lernzuwachs. Die Schülerinnen und Schüler können also selbst gut einschätzen, ob sie „etwas lernen“: Je stärker die Klasse über zu hohes Unterrichtstempo klagt, desto schwächer ist – unter Berücksichtigung der Ausgangsbedingungen – der tatsächliche Kompetenzzuwachs. Es lohnt sich also für Deutschlehrer, sich mit der Unterrichtswahrnehmung ihrer Schülerinnen und Schüler auseinanderzusetzen.

In Übereinstimmung mit einem häufigen Befund der Schulforschung zeigt sich auch ein klarer Zusammenhang zwischen Leistungsentwicklung und Anspruchsniveau. In leistungsstärkeren Klassen nehmen die Schülerinnen und Schüler wahr, dass die Lehrkraft hohe Ansprüche stellt, aber ein hohes Anspruchsniveau hat auch seinerseits – unter Kontrolle der Lernvoraussetzungen und weiterer Hintergrundmerkmale – einen signifikanten positiven Effekt auf die Kompetenzentwicklung (zur Diskussion um die Stärke der Effekte vgl. Kap. 26).

Ebenfalls erwartungsgemäß zeigt sich die Disziplin im Unterricht als hochrelevante Komponente der Unterrichtsqualität: In Klassen, deren Leistungen angesichts ihrer Rahmenbedingungen besonders positiv sind, berichten die Schülerinnen und Schüler von einer positiven Disziplin, und diese wirkt sich zumindest im Bereich Sprachbewusstheit (d.h. bei demjenigen Deutschtest, der vergleichsweise änderungs-sensitiv ist) schwach signifikant auf den Lernzuwachs aus.

Hinsichtlich einer ganzen Reihe von Unterrichtsmerkmalen, die häufig mit „guten Unterricht“ oder gar Leistungszuwachs assoziiert werden, zeigt sich jedoch ein anderes, zunächst überraschendes Muster: Je schlechter die Ausgangsleistung einer Klasse ist, umso strukturierter und unterstützender wird unterrichtet, und desto mehr Lerngelegenheiten – sei es bezüglich Wortschatz, Sprachbewusstheit, Rechtschreibung oder Briefe schreiben – werden angeboten. Offensichtlich reagieren Deutschlehrkräfte in diesem Sinne adaptiv auf problematische Lernausgangslagen. Interessanterweise lässt sich aber für keines dieser Unterrichtsmerkmale eine Wirksamkeit im Sinne eines nachweisbaren Effekts auf die weitere Kompetenzentwicklung nachweisen.

Diese Analyse wird bestätigt auf der Basis der Lehrereinschätzungen: In Klassen, die vergleichsweise schlechte Ausgangsbedingungen aufweisen, berichten auch die unterrichtenden Lehrer von größerer Methodenvielfalt unter Einschluss alternativer Unterrichtsmethoden sowie von häufiger Verwendung rechtschreib- und grammatikförderlicher Unterrichtselemente. Statistisch lässt sich jedoch für diese offenbar remedialen Maßnahmen kein leistungsförderlicher Effekt nachweisen.

Nur in wenigen Fällen lässt sich ein positiver Einfluss auf die Leistungsentwicklung belegen: Schülerorientierung sowie die intensive Nutzung von Hausaufgaben gehen mit einem Zuwachs an Kompetenz im Bereich Sprachbewusstheit einher, während hohe Verständlichkeit und eine Betonung von Prosa-Texten im Unterricht schwach, aber signifikant positiv die Lesekompetenz fördern.

Im Kontrast zu den verschiedenen Unterrichtsmerkmalen, die offensichtlich remedial in schwachen Klassen eingesetzt werden, zeigen sich einige wenige Merkmale, die bevorzugt in leistungsstärkeren Klassen auftreten: Gruppenarbeit sowie (dem Urteil der Lehrkräfte zufolge) ein stärker literarischer Unterricht.

Der remediale Charakter vieler Unterrichtsmethoden korrespondiert mit analogen Befunden auf der Schulebene (vgl. Kap. 34): Auch dort werden bestimmte Arten der Lehrerkooperation und anderes mehr als Reaktion auf schwache Leistungen erkannt, deren Wirksamkeit im Sinne einer Leistungssteigerung jedoch nicht belegt werden kann. Möglicherweise stellt die Schul- und Unterrichtspraxis in Deutschland (zu) sehr darauf ab, mit einer Reihe von Maßnahmen und Methoden schwache Klassen zu fördern, offenbar ohne ein angemessenes Anspruchsniveau, das ein bedeutsamer Wirkungsfaktor zu sein scheint.

Je wichtiger die Lehrkräfte – aus der Perspektive ihrer Schüler gesehen – sprachliche Basiskompetenzen wie richtiges Sprechen und Schreiben oder angemessene Wortwahl nehmen, desto stärker fällt der Leistungszuwachs bei DESI aus. Damit ergibt sich ein eindeutiges Plädoyer für einen sprachbewussten, in diesem Sinne kognitiv aktivierenden Unterricht.

Die Wirksamkeit von Faktoren wie Anspruchsniveau, Verständlichkeit und Unterrichtsdisziplin stimmt mit Grundannahmen der Unterrichtsqualitätsforschung überein (vgl. etwa Klieme 2006). Aus diesen theoretischen Bezügen lässt sich auch ableiten, dass Merkmale wie die Vielfalt von Lernsituationen, Unterstützung und Unterrichtsklima eher die Lernmotivation als die kognitiven Kompetenzen fördern. Entsprechende Analysen werden in weiterführenden Publikationen vorgelegt.

Literatur

- Abraham, U./Bremerich-Vos, A./Frederking, V./Wieler, P. (Hrsg.) (2003): Deutschdidaktik und Deutschunterricht nach PISA. Freiburg: Fillibach.
- Baumert, J./Trautwein, U./Artelt, C. (2003): Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In: Baumert, J./Artelt, C./Klieme, E. (Hrsg.): PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland, S. 261-332. Opladen, Leske+Budrich.
- Baumert, J./Kunter, M./Brunner, M./Krauss, S./Blum, W./Neubrand, M. (2004): Schule und Unterricht. Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In: Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M./Pekrun, R./Rolff, H.-G./Rost, J./Schiefele, U. (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann, S. 314-349.
- Bremerich-Vos, A. (2006): Deutsch. In: Arnold, K. H./Sandfuß, U./Wichmann, J. (Hrsg.): Handbuch Unterricht. Bad Heilbronn: Klinkhardt. S. 449-453.
- Clausen, M. (2002): Unterrichtsqualität: Eine Frage der Perspektive? Münster: Waxmann.
- Dehn, M./Hüttis.Graff, P. (2005): Kompetenz und Leistung im Deutschunterricht. Freiburg: Fillibach.
- Ditton, H. (1998): Mehrebenenanalyse. Grundlagen und Anwendungen des Hierarchisch Linearen Modells. Weinheim, Juventa.
- Eikenbusch, G. (2001): Qualität im Deutschunterricht der Sekundarstufe I und II. Berlin: Cornelsen.
- Frey, H. (2007): Kann eine Vermittlung von Lesestrategien die Lesekompetenz verbessern? In: Willenberg, H. (Hrsg.): Kompetenzhandbuch für den Deutschunterricht. Baltmannsweiler: Schneider Verlag Hohengehren, S. 188-198.

- Fritzsche, J. (1994a): Zur Didaktik und Methodik des Deutschunterrichts. 2. Schriftliches Arbeiten. Stuttgart: Klett.
- Fritzsche, J. (1994b): Zur Didaktik und Methodik des Deutschunterrichts. 3. Umgang mit Literatur. Stuttgart: Klett.
- Helmke, A./Schrader, F.-W. (2001): Jenseits von TIMMS: Messungen sprachlicher Kompetenzen, komplexe Längsschnittstudien und kulturvergleichende Analysen. In: Weinert, F. (Hrsg.): Leistungsmessungen in Schulen, S. 237-250. Weinheim: Beltz.
- Hosenfeld, I./Helmke, A./Schrader, F.-W. (2002): Zur Rolle des Klassenkontextes für Geschlechtsunterschiede bei Mathematikleistungen – Ergebnisse der MARKUS-Studie. Beitrag präsentiert bei 43. Kongress der Deutschen Gesellschaft für Psychologie, Berlin.
- Hox, J.J. (2002): *Multilevel Analysis. Techniques and Applications*. Mahwah, New Jersey, Lawrence Erlbaum Associates.
- Kämper-van den Boogaart (Hrsg.) (2003): *Deutsch Didaktik Leitfaden für die Sekundarstufe I und II*. Berlin: Cornelsen.
- Klieme, E. (2006): Empirische Unterrichtsforschung: Aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. In: Zeitschrift für Pädagogik, 52, Heft 6, S. 765-773.
- Klieme, E./Lipowsky, F./Rakoczy, K./Ratzka, N. (2006): Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In: Prenzel, M./Allolio-Näcke, L. (Hrsg.): Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms, S. 128-146.
- Klieme, E./Rakoczy, K. (2003): Unterrichtsqualität aus Schülerperspektive: Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht. In: Deutsches PISA – Konsortium (Hrsg.): PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen: Leske + Budrich, S. 333-360.
- Köller, O./Trautwein, U. (2004): Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In Köller, O./Watermann, R./Trautwein, U./Lüdtke, O. (Hrsg.): Wege zur Hochschulreife in Baden-Württemberg, S. 285-326. Opladen: Leske+Budrich.
- Lipowsky, F./Rakoczy, K./Klieme, E./Reusser, K./Pauli, C. (2004). Hausaufgabenpraxis im Mathematikunterricht – ein Thema für die Unterrichtsqualitätsforschung? In Doll, J./Prenzel, P. (Hrsg.). Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung, S. 250-266. Münster: Waxmann.
- Muthén, B. (2002): Beyond SEM: General latent variable modeling. *Behaviormetrika* 29(1): S. 81-117.
- Rakoczy, K./Klieme, E./Drollinger-Vetter, B./Lipowsky, F./Pauli, C./Reusser, K. (2007). *Structure as a Quality Feature in Mathematics Instruction: Cognitive and Motivational Effects of a Structured Organisation of the Learning Environment vs. a Structured Presentation of Learning Content*. In Prenzel, M. (Hrsg.). *Studies on the educational quality of schools. The final report on the DFG Priority Programme*, S. 102-121. Münster: Waxmann.
- Sawyer, W./Van den Ven, P.-H. (2006): Starting points. Paradigms in Mother tongue Education. In: L1 – Educational Studies in Language and Literature, 7 (1), pp. 5-20.
- Trautwein, U./Köller, O./Baumert, J. (2001): Lieber oft als zu viel: Hausaufgaben und die Entwicklung von Leistung und Interesse im Mathematik-Unterricht der 7. Jahrgangsstufe. *Zeitschrift für Pädagogik*, 47(5), 703-724
- Willenberg, H. (Hrsg.) (2007): *Kompetenzhandbuch für den Deutschunterricht*. Baltmannsweiler: Schneider Verlag Hohengehren.

Tuyet Helmke / Andreas Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Günter Nold / Konrad Schröder

29 Die Videostudie des Englischunterrichts

In der DESI-Hauptuntersuchung wurde der Unterricht sehr differenziert aus Sicht der Schülerinnen und Schüler erfasst. Die über Fragebogen erhobenen und auf den jeweiligen Unterricht in den Klassen bezogenen Schülerangaben sind eine aussagekräftige Informationsquelle, da sie sich auf einen längeren Beobachtungszeitraum beziehen. Um darüber hinaus detaillierte Einblicke in das Unterrichtsgeschehen zu erhalten, sind direkte Beobachtungen des Lehrer- und Schülerverhaltens bzw. der Lehrer-Schüler-Interaktion unverzichtbar. Dies geschieht mittlerweile mit Hilfe von Videoaufzeichnungen. Die DESI-Videostudie verfolgt drei zentrale Ziele:

- Beschreibung und Analyse der realen Unterrichtspraxis, insbesondere bei Themen, die sich einer Befragung von Lehrkräften oder Schülern entziehen oder für die ein mehrperspektivischer Ansatz sinnvoll erscheint;
- Gewinnung verhaltensnaher Indikatoren der Unterrichtsqualität, um so genauere Hinweise zu deren Grundlagen und ihrer Rolle für die sprachliche Entwicklung zu erhalten;
- Bestandsaufnahme der mündlichen Kommunikation von Schülerinnen und Schülern.

Um diese Ziele zu erreichen, um eine möglichst hohe Vergleichbarkeit der Unterrichtsvideografie zu gewährleisten und um deren Akzeptanz nicht unnötig einzuschränken, wurden die folgenden Vorgaben gemacht:

- Pro Klasse wurden zwei *Unterrichtsstunden* aufgenommen, die zwei unterschiedliche Themenschwerpunkte aufweisen sollten: ein stärker *sprachlernorientiertes* Thema und ein *interkulturelles* Thema. In diesem Kapitel differenzieren wir nicht zwischen diesen beiden Stunden.
- Die inhaltliche und methodische Ausgestaltung dieser beiden Stunden blieb der Lehrperson völlig überlassen. Es wurde lediglich darauf hingewiesen, dass Schüler oft zum *freien Sprechen* kommen sollten, deswegen seien Inhalte wünschenswert, zu denen möglichst *alle Schüler* der Klasse etwas beitragen können.
- Da es aus technischen Gründen nicht möglich war, Sprachäußerungen von *Kleingruppen* simultan aufzuzeichnen, sollte diese Lehrform (das gleiche gilt für längere Stillarbeitsphasen etc., die für die Videografie nicht ergiebig sind) zumindest nicht im Vordergrund stehen.

Die Grundlage für die Konzipierung der Videostudie und ihrer Instrumente sind Theorien und empirische Ergebnisse der (allerdings kaum auf Fremdsprachenunterricht bezogenen) Lehr-Lern-Forschung (Helmke 2004; Wellenreuther 2004) sowie aus dem US-amerikanischen Raum stammende Darstellungen des Forschungsstandes zum Fremdsprachenunterricht (Waxman/Hilberg/Tharp 2004; Spada/Fröhlich 1995; Judd/Tan/Walberg 2001). Dazu kommen deutsche Lehrbücher und

Überblicksdarstellungen der Englischdidaktik (Heuer/Klippel 1993; Timm 1998; Bausch/Christ/Krumm 2002). Daneben ließen wir uns von den Erfahrungen, Instrumenten und Ergebnissen der großen internationalen Videostudien des Unterrichts leiten (Baumert u.a. 1997; Stigler/Gallimore/Hiebert 2000; Reusser/Pauli 2004; Hiebert u.a. 2003; Prenzel u.a. 2002; Klieme/Reusser 2003).

29.1 Stichprobe und Methode

An der Videostudie¹ nahmen insgesamt 105 Klassen teil, davon 15 Klassen im Bildungsgang Hauptschule (14%), 33 Klassen in der Realschule (31%), 53 Klassen im Gymnasium (51%) und 4 Klassen in der Integrierten Gesamtschule (4%). Der Prozentanteil aller Bildungsgänge bei der DESI-Videografie im Vergleich mit der gesamten DESI-Stichprobe ist in Tabelle 29.1 dargestellt.

Tabelle 29.1: Verteilung der Klassen nach Bildungsgängen in der DESI-Videostudie.

Bildungsgang	Anzahl (Videostudie)	Prozent (Videostudie)	Prozent (DESI insg.)
Hauptschule	15	14	22
Realschule	33	31	32
Gymnasium	53	51	38
Integrierte Gesamtschule	4	4	8

Aus Tabelle 29.1 geht hervor, dass die Hauptschule etwas unter- und das Gymnasium etwas überrepräsentiert ist. Für die Verallgemeinerbarkeit der Befunde ist dies jedoch nicht von Bedeutung, da für die Analyse der videobasierten Ergebnisse spezielle Gewichte verwendet werden.

Zur Überprüfung der Repräsentativität der Videoklassen durchgeführte Analysen zeigen, dass sich die beiden Gruppen nur hinsichtlich weniger Variablen signifikant voneinander unterscheiden: Insbesondere unterscheidet sich die an der Videostudie teilnehmende Schülergruppe von der Vergleichsgruppe *nicht* signifikant hinsichtlich der Englischtestleistung, wie Tabelle 29.2 zeigt: Die Gymnasiallehrpersonen der Videostudie sind im Durchschnitt etwas jünger (und demnach weniger berufserfahren) und legen etwas mehr Wert auf Leistungsdifferenzierung, und die Hauptschullehrer der Videostudie haben mehr Kontakte zum englischsprachigen Ausland als die Vergleichsgruppe. Insgesamt gesehen sind die Unterschiede so geringfügig, dass die Verallgemeinerbarkeit der Ergebnisse nicht beeinträchtigt ist.

¹ Planung, Durchführung, Datenerhebung und -analyse sowie videobasierte Ergebnisrückmeldung an die Lehrkräfte lagen in der Hand der Landauer Arbeitsgruppe (A. Helmke, T. Helmke, F.-W. Schrader und W. Wagner).

Tabelle 29.2: Mittlere Englisch-Sprachkompetenzen der an der Videostudie teilnehmenden und nicht teilnehmenden Schülerinnen und Schüler.

Bildungsgang Teilnahme an Videostudie	Hauptschule		Realschule		Gymnasium		IGS	
	ja	nein	ja	nein	ja	nein	ja	nein
Englischleistung (Mittelwert)	424	413	502	500	615	610	461	420
Standardfehler p (Teilnahme vs. Nicht-Teilnahme)	8.29	3.27	8.27	5.59	7.99	4.95	47.81	13.20
	.16		.86		.56		.38	

29.2 Instrumente der Videostudie

Die den Analysen zugrunde liegenden Daten wurden mittels (a) einer mikroanalytischen Basiskodierung des videografierten Unterrichts, (b) eines stundenbezogenen Ratings, (c) eines Schüler- und (d) eines Lehrerkurzfragebogens gewonnen².

Basiskodierung

Die Auswertung der digitalisierten videografierten Unterrichtsstunden erfolgte mit Hilfe des Programms *Videograph* (Rimmele 2005) in zwei Schritten:

(1) *Transkription* des Unterrichtsgeschehens, insbesondere aller sprachlichen Äußerungen, als Voraussetzung für die weiteren Auswertungen. Eines der Ergebnisse der Video-Pilotierung war angesichts oft sehr langer, aus mehreren Sätzen mit mehreren Aspekten wie z.B. Rückmeldung an Schüler/in, der/die gerade „dran“ ist, und neue Frage an Klasse, bestehender Lehreräußerungen eine *satzweise* Kodierung (Grundlage: ein einzelner Satz – anstelle der verbreiteten Analyseeinheit *Turn* im Sinne von Sprecherwechsel).

(2) *Kodierung*: Alle Einzeläußerungen bzw. nicht-sprachlichen Interaktionen wurden – wie bei TIMSS-Video – lückenlos kodiert („Basiskodierung“), so dass sich ein vollständiges Bild des Unterrichtsgeschehens und der Lehrer-Schüler-Interaktion ergibt. Alle Kodierungen können zeit- oder häufigkeitsbasiert ausgewertet werden, also z.B. die Anzahl von Lehrerfragen pro Stunde und/oder der Zeitanteil der Stunde, die auf Lehrerfragen entfällt. Entsprechend den Zielen der DESI-Videostudie wurden die zentralen Kategorien des bewährten Verfahrens COLT (Communicative Orientation of Language Teaching, Spada/Fröhlich 1995) für die Basiskodierung adaptiert und durch eigene Konstrukte ergänzt. Die wichtigsten Kategorien, deren Vorkommenshäufigkeit und zeitliche Dauer erfasst wurden, sind:

- Interaktionsteilnehmer (Wer spricht zu wem?)
- Sprechanteile Lehrer / Schüler; nonverbale Phase

2 Eine ausführliche Beschreibung der Logistik, Technik, Instrumente, Kodierschemata und Verfahren der Datenanalyse sowie der Ergebnissrückmeldung an die beteiligten Lehrkräfte findet sich in einem von den Autoren des vorliegenden Kapitels herausgegebenen *Technischen Bericht*.

- Benutzte Sprache (Englisch, Deutsch, gemischt)
- Gegenstand der Äußerung (lehrstoffbezogen, disziplinbezogen, prozedural, sozial)
- Länge der Schüleräußerung (Ein-Wort-Äußerung, Satzfragment, ganzer Satz, Satzunterbrechung durch Lehrer)
- Art der Schüleräußerung (Wiederholen, Ablesen, nach Vorgabe sprechen, frei sprechen)
- Korrektheit der Schüleräußerung
- Funktion der Lehreraktivität (Darstellung, strukturierender Hinweis, Frage, Anweisung, Aufforderung)
- Funktion der Lehrerreaktion (Hilfestellung, Rückmeldung, Fehlerbehandlung, Selbstbeantwortung)
- Charakteristika der Frage (Antwortspielraum, sprachliche Komplexität der vom Schüler geforderten Antwort, Authentizität, Lebensweltbezug)
- Gehalt der Lehrerrückmeldung (informativ und affektiv)
- Umgang mit Fehlern (wie, wann und von wem werden welche Arten von Fehlern korrigiert?)

Die Kategorien lassen sich *kontextuieren* (d.h. je nach Unterrichtskontext oder Episode aufschlüsseln), *kombinieren* (z.B. Anteil von Satzfragmenten in deutscher vs. englischer Sprache) und *sequenzieren* (Analyse von Mustern der zeitlichen Aufeinanderfolge: Ketten, Skripts, Sequenzen – z.B. Häufigkeit von Lehrerfragen, denen eine Wartezeit folgt, oder Häufigkeit, mit der ein- und derselbe Schüler in einen Lehrer-Schüler-Dialog einbezogen wird).

Rating

Zusätzlich zu den niedrig-inferenten Beobachtungskategorien der Basiskodierung wurden hochinferente Beurteilungen (Ratings) wichtiger Aspekte der Unterrichtsqualität³ auf der Ebene der gesamten Stunde vorgenommen. Basis sind die klassischen Ratingbögen zur Unterrichtsqualität, wie sie beispielsweise bei der Deutschen Classroom Environment Study (Helmke/Schrader 1993) oder bei der Grundschulstudie SCHOLASTIK verwendet wurden (Helmke/Schrader 1997), ergänzt um Aspekte, die für den Fremdsprachenunterricht spezifisch sind:

- *Einhaltung der DESI-Vorgaben* (als Kontrolle): Spracharbeit; interkultureller Bezug; Einbezug möglichst aller Schüler
- *Zielorientierung*: Kommunikation; Anregung von Lern- und Denkstrategien; Eingehen auf Lern- und Verständnisschwierigkeiten; Anregung anspruchsvoller, weiterführender Themen

3 Niedrig- bzw. hoch-inferent: niedriger bzw. relativ hoher Ermessens- und Beurteilungsspielraum. Um die Qualität des Ratings zu sichern, wurden regelmäßig Reliabilitätskontrollen durchgeführt. Als Maß für die Beobachterübereinstimmung wird die Intraklassen-Korrelation (ICC = Intraclass Correlation) verwendet (s. Wirtz/Casparly 2002). Hohe Werte geben an, dass Unterschiede zwischen den Beobachtungseinheiten (hier: den einzelnen Videofilmen) einen geringen Fehleranteil (d.h. geringe Unterschiede zwischen den Urteilern) aufweisen und damit zuverlässig beurteilt werden.

- *Flüssigkeit vs. Genauigkeit*: Forcierung von Korrektheit; Wertschätzung von Flüssigkeit
- *Klarheit*: Kohärenz (Inhaltsaspekt); Prägnanz der Lehrersprache (Sprachaspekt)
- *Strukturiertheit*: Systematischer, sachlogisch geordneter Stundenaufbau; Previews, Zusammenfassungen, Reviews; Hervorhebungen und Hinweise; Aktivierung oder Schaffung von Vorwissen
- *Lenkung*: Engführung; Schülerorientierung
- *Klassenführung*: Aufgabenorientierung; Unterrichtsplanung und Zeitmanagement; Störungsfreiheit
- *Unterrichtsklima*: Wärme, Herzlichkeit; Humor
- *Fehlerklima*: Positiver vs. negativer Umgang mit Schülerfehlern seitens der Lehrperson und seitens der Klasse
- *Motivierungsqualität*: Lehrerengagement; Schülerengagement; Lebensweltbezug und Authentizität; Ermutigung von Schüleräußerungen
- *Passung*: Variation des Unterrichts/Adaptivität; Angemessenheit des Anforderungsniveaus
- *Lehrersprache*: Sprechweise; Aussprache; Wortschatzsicherheit; Grammatik-sicherheit; inhaltliche Korrektheit; verbale Tics und Marotten

Alle Ratings sowie die Antwortkategorien (*trifft nicht zu / trifft eher nicht zu / trifft eher zu / trifft zu*) werden für die Rater/innen anhand von Beispielen detailliert beschrieben.

Episodenbildung

Episoden sind nach didaktischer Funktion, Arbeits- und Sozialform o.ä. mehr oder weniger klar voneinander abgrenzbare Phasen im Unterricht. Der Grund für eine Klassifikation unterhalb der Ebene der gesamten Stunde ist der, dass Unterrichtsmaßnahmen oder Lehrerverhaltensweisen je nach Episode eine unterschiedliche Bedeutung haben können. Für die DESI-Videostudie wurden die folgenden Episoden (Grundlage: Arbeits- und Sozialform) verwendet: Lehrervortrag; lehrerzentrierte Gesprächsführung; Schülervortrag; schülerzentrierte Gesprächsführung; Lernspiel; Verwendung von audiovisuellen Materialien; Einzelarbeit; Partnerarbeit; Gruppenarbeit; Übergang.

Schüler- und Lehrerkurzfragebogen

Um mehr über die Lehr-Lern-Prozesse und deren Bedingungen während der videografierten Unterrichtsstunde zu erfahren, wurden nach dem Unterrichtsende ein Lehrer- und ein Schülerkurzfragebogen eingesetzt. Im *Schülerkurzfragebogen* geht es um: Verständlichkeit des Unterrichtsstoffs, Aufmerksamkeit, subjektive Schwierigkeit, Interessantheit, aktive Beteiligung am Unterricht, Repräsentativität des videografierten Unterrichts. Der *Lehrerkurzfragebogen* bezieht sich auf die fol-

genden Aspekte: Nervosität der Lehrkraft während der Aufzeichnung; zentrales Thema der aufgezeichneten Stunde; Zufriedenheit mit der aufgezeichneten Stunde; Zufriedenheit mit dem Verhalten der Klasse; geplanter Ablauf der Stunde und eventuelle Abweichungen; Einschätzungen des Verhaltens der Schülerinnen und Schüler (Aufmerksamkeit, Über-/Unterforderung etc.); Einschätzung des eigenen Verhaltens (Sprechzeit, Häufigkeit der gestellten und beantworteten Fragen, der vorgekommenen und korrigierten Fehler; Repräsentativität der aufgezeichneten Stunde).

29.3 Ausgewählte deskriptive Ergebnisse

Im Folgenden werden ausgewählte Ergebnisse der Video-Basiskodierung dargestellt.

Unterrichtssprache und Sprechanteile

In Abbildung 29.1 sind die Sprechanteile der Lehrkraft sowie der Schülerinnen und Schüler nach der verwendeten Sprache aufgeschlüsselt: Englisch, Deutsch (D), gemischt („code switching“) (G) sowie „unverständlich“ (U) bzw. „nicht zuzuordnen“ (N) (dazu gehört bei den Lehrerinnen und Lehrern z.B. das Nennen von Schülernamen).

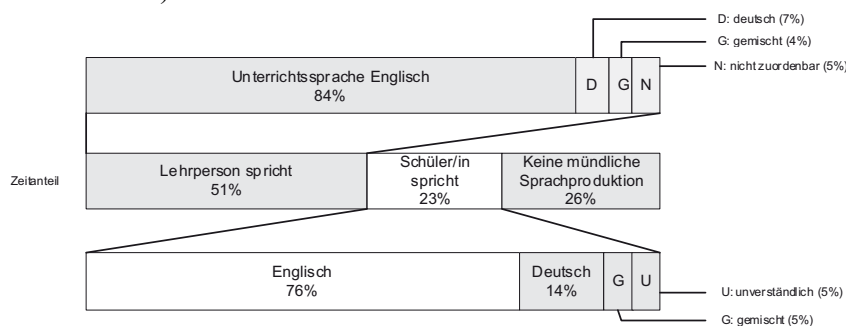


Abbildung 29.1: Unterrichtssprache und Sprechanteile von Lehrern und Schülern im Englischunterricht.

Der obere Balken zeigt, dass der größte Teil (84%) aller im Unterricht vorfindbaren Lehreräußerungen auf Englisch erfolgt. Die Schülerinnen und Schüler sprechen in 76% der Unterrichtszeit Englisch, wie der untere Balken zeigt. Der mittlere Balken zeigt, dass Lehrerinnen und Lehrer im Durchschnitt mehr als doppelt so lange sprechen wie alle Schülerinnen und Schüler zusammen. Das restliche Viertel der Zeit („keine mündliche Sprachproduktion“) umfasst Aktivitäten wie Still- oder Einzelarbeit, Übergänge, Vorbereitung von Medien oder Wartezeit.

Die Schulklassenunterschiede in den Sprechzeiten (prozentualer Anteil von Schülersprechanteilen an der Gesamtsprechzeit) reichen von minimal 12% bis maximal 71%. Der mittlere Sprechanteil der Schülerinnen und Schüler (bezogen auf die gesamte Sprechzeit, nicht wie in Abb. 29.1 auf die gesamte Unterrichtszeit) in einer

Unterrichtsstunde liegt bei 32% (Hauptschule: 35%, Realschule: 27%, Integrierte Gesamtschule: 25%, Gymnasium: 37%). Verglichen mit der in der TIMSS-Videostudie (Klieme/Schümer/Knoll 2001) für das Fach Mathematik in der achten Jahrgangsstufe gefundenen Lehrerdominanz (Lehrer-Zeitanteil an öffentlicher Rede: 76%) ist der Sprechanteil der Lehrerinnen und Lehrer im Englischunterricht (68%) niedriger – die Schülerinnen und Schüler kommen häufiger zu Wort.

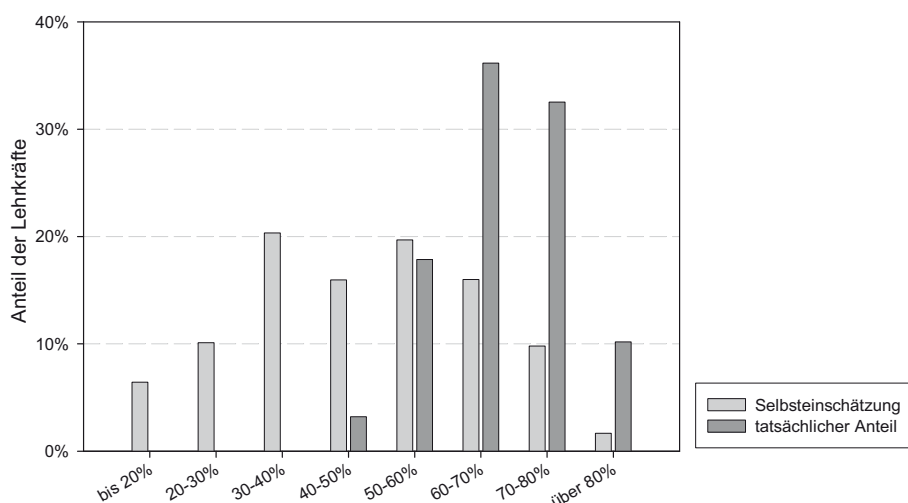
Der geringe Schüleranteil an der Sprechzeit ist möglicherweise dann kein Problem, wenn der classroom discourse des Lehrers authentisch, reich und vielfältig ist (*comprehensible input* im Sinne von Krashen, vgl. Krashen/Terrell 1983)⁴. Bei einer verarmten, stereotypen Lehrersprache könnte er sich aber ungünstig auswirken. Wenn (wie dies der Fall ist) mehr als die Hälfte der Sprechzeit des Lehrers für Hilfestellungen und Aufforderungen und nur ein Viertel für Fragen und Rückmeldungen aufgewendet wird, dann kann der sprachliche Input des Lehrers nur bedingt als Lernanreiz für die Schülerschaft angesehen werden. Hinzu kommt, dass die Lehrerfragen zu fast 70% in die Kategorie „niedrige Authentizität“ und zu mehr als 80% in die Kategorie „niedriger Lebensweltbezug“ fallen.

Genauigkeit der Selbsteinschätzung

Eine für die Unterrichtsentwicklung wichtige Frage ist, in welchem Maße sich Lehrkräfte eigentlich wichtiger Merkmale ihres eigenen Unterrichts bewusst sind. Bei DESI wurden die Lehrpersonen deshalb unmittelbar nach der videografierten Unterrichtsstunde u.a. danach gefragt, welcher Anteil der gesamten Sprechzeit in etwa auf das eigene Reden entfiel. Abbildung 29.2 zeigt, dass Lehrkräfte insgesamt gesehen, d.h. über alle Klassen hinweg, ihre eigene Sprechzeit unterschätzen. Anders ausgedrückt: Es zeigt sich eine Kluft zwischen subjektiver Einschätzung (episodisches Gedächtnis) und der videobasierten Realität (Echtzeit).

Laut Videoaufzeichnung beträgt sie im Durchschnitt 68%, laut Selbsteinschätzung 51%. Die Verteilung der Balken in Abbildung 29.2 macht dies auch auf andere Weise deutlich: Etwa ein Drittel der Lehrkräfte schätzt den eigenen Anteil an der Sprechzeit als gering ein (höchstens 40%) – solch geringe Sprechquoten kommen jedoch laut Videostudie in keinem einzigen Fall vor, wie aus dem Fehlen der schwarzen Balken zu ersehen ist.

4 Dies wird in weiteren differenzierten videobasierten Analysen der Lehrersprache und des classroom discourse erfolgen.



Sprechanteil der Lehrkraft an der Gesamtsprechzeit

Abbildung 29.2: Geschätzte und reale Lehrersprechzeit im Englischunterricht.

Arten sprachlicher Schüleräußerungen

Schlüsselt man den Sprechanteil der Schülerinnen und Schüler nach der Art der Äußerungen auf (vgl. Abbildung 29.3), so erkennt man, dass in knapp der Hälfte dieser Sprechzeit in englischer Sprache frei gesprochen wird, ein gutes Viertel auf Ablesen (z.B. Vorlesen) und ein Fünftel auf andere Sprechaktivitäten wie Wiederholen oder Nachsprechen entfällt.

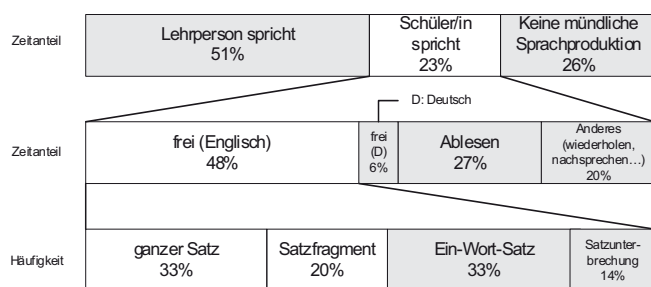


Abbildung 29.3: Art und Länge von Schüleräußerungen im Englischunterricht.

Freies Sprechen wurde kodiert, wenn eine Äußerung von der Schülerin oder vom Schüler selbstständig formuliert wird und diese(r) bei der Beantwortung nicht an Vorgaben gebunden ist. *Ablesen* kann sich sowohl auf eigene als auch auf fremde Texte beziehen. Beim freien Sprechen wird die deutsche Sprache nur zu einem geringen Teil benutzt (beim Ablesen und anderen Schüleräußerungen ist sie nicht separat ausgewiesen). Schlüsselt man das freie Sprechen nach Länge bzw. Vollständigkeit

der Äußerungen näher auf, so zeigt sich, dass jeweils etwa ein Drittel davon auf ganze Sätze und auf Ein-Wort-Äußerungen fällt. Der Rest sind unvollständige Sätze, die entweder von der Schülerin oder vom Schüler selbst in dieser Form produziert werden (Satzfragmente, z.B. Aufzählungen oder andere nicht in einen ganzen Satz eingekleidete Äußerungen) oder durch Unterbrechungen seitens der Lehrkraft oder anderer Schülerinnen und Schüler zustande kommen.

Im Bereich der Schüleräußerungen fällt auf, dass ein relativ hoher zeitlicher Anteil mit Ablesen, Wiederholen, Nachsprechen und dergleichen verbracht wird (47%). Dies ist nicht per se negativ zu bewerten, doch es stellt sich heraus, dass nur rund 11% der gesamten Unterrichtszeit auf frei formulierte Schüleräußerungen in englischer Sprache entfällt. Bei einem Drittel der frei formulierten Antworten handelt es sich zudem um „Ein-Wort-Äußerungen“. Die Analyse zeigt insgesamt, dass der sprachlichen Produktion durch die Lernenden in einem im Wesentlichen einsprachigen Englischunterricht enge Grenzen gesetzt sind.

Umgang mit Schülerfehlern

Abbildung 29.4 knüpft ebenfalls an die zuvor berichteten Sprechanteile von Schülerinnen und Schülern an und schlüsselt die Schüleräußerungen danach auf, ob diese richtig oder fehlerhaft sind und wie dann mit fehlerhaften Äußerungen umgegangen wird.

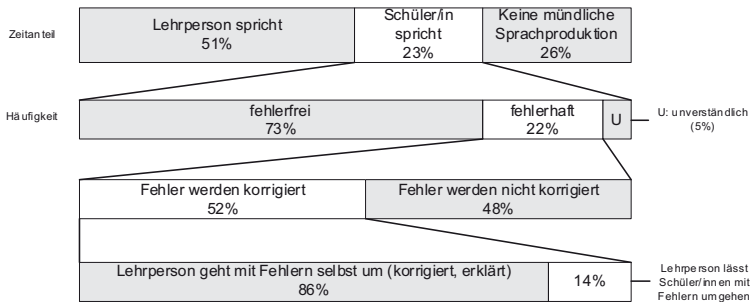


Abbildung 29.4: Umgang mit Schülerfehlern im Englischunterricht.

Ein gutes Fünftel aller Schüleräußerungen ist fehlerhaft. In gut der Hälfte der fehlerhaften Schüleräußerungen wird der Fehler korrigiert, wobei nur in knapp 15% der Fälle die Schülerinnen und Schüler die Gelegenheit erhalten, dies selbst zu tun („Selbstkorrektur“). In der überwiegenden Mehrzahl der Fälle korrigieren die Lehrerin oder der Lehrer.

Die Ergebnisse sprechen dafür, dass eine durchgängige, strikte Fehlerkorrektur im heutigen Englischunterricht nicht stattfindet. Ob dies bewusst intendiert ist und somit fachdidaktischen Grundsätzen entspricht, geht aus den erhobenen Daten nicht hervor. Peer-Korrektur als die aus fachdidaktischer Sicht günstigste Art der Behandlung von Fehlern kommt im Unterricht kaum vor: In mehr als zwei Drittel aller Korrekturfälle korrigiert die Lehrkraft die Fehler selbst, gewissermaßen en passant, denn in lediglich (weiteren) 8% der Fälle erfolgt eine Erklärung des Fehlers.

Die Schüler werden nur in 14% aller Korrekturfälle produktiv-kognitiv involviert. In 10% aller Korrekturfälle erfolgt – als Anreiz zur Selbstkorrektur durch die Schüler – ein allgemeinerer Hinweis auf den Fehler durch die Lehrkraft.

Wartezeit nach Lehrerfragen

Wie reagieren Lehrpersonen, wenn ihre Fragen nicht spontan (d.h. innerhalb eines Toleranzintervalls von drei Sekunden) beantwortet werden? Dies zeigt Abbildung 29.5.

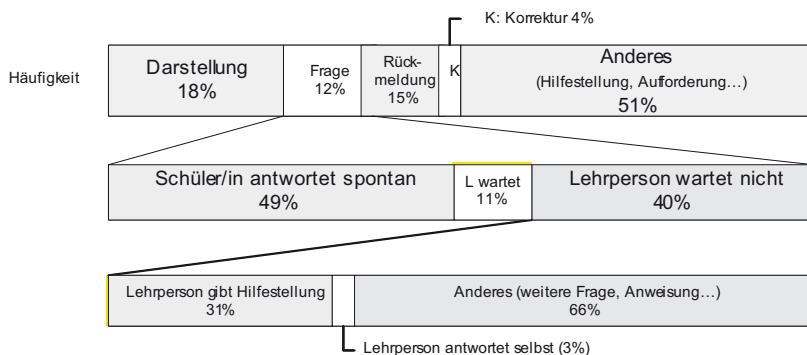


Abbildung 29.5: Wartezeit nach Lehrerfragen im Englischunterricht.

Knapp die Hälfte der Lehrerfragen wird von den Schülerinnen und Schülern kurzfristig beantwortet. Ist dies nicht der Fall, d.h. vergehen mehr als drei Sekunden, haben Lehrpersonen verschiedene Möglichkeiten zu reagieren: Hilfestellung geben, eine andere Frage stellen, die Frage weiterreichen, im Stoff weitergehen etc. Dies tun Lehrkräfte in ca. 40% der Fälle. Sie können jedoch auch *warten*, also der gefragten Schülerin oder dem gefragten Schüler Zeit geben, eine Antwort zu überlegen. Dies passiert allerdings nur in ca. 11% der Fälle. Nimmt man die Wartezeit unter die Lupe, dann ergibt sich, dass sie von drei Sekunden (definitionsgemäß die untere Grenze) bis knapp zehn Sekunden reicht. Der Mittelwert liegt bei 6.7 Sekunden.

Dialoge

In aller Regel ist die sprachliche Lehrer-Schüler-Kommunikation (Ausgangspunkt: eine Lehrerfrage) dadurch charakterisiert, dass nach erfolgter Antwort eine andere Schülerin oder ein anderer Schüler „drankommt“. Die Datenstruktur der DESI-Videostudie gestattet es aber auch, Ketten von Lehrer-Schüler-Interaktionen mit *derselben* Schülerin bzw. *demselden* Schüler zu identifizieren, d.h. Sequenzen des Lehrer-Schüler-DIALOGS, die mehr als drei Stationen aufweisen (vgl. Abbildung 29.6).

Vorausgegangen ist dem in dieser Abbildung gezeigten Muster die Standardsituation: Lehrerfrage (Station 1), Schülerantwort (Station 2) und Lehrerreaktion, z.B. Feedback (Station 3). Erwartungsgemäß nimmt die Häufigkeit von längeren Dialogen nach der dritten Station erheblich ab. Es finden sich aber gelegentlich auch längere

Lehrer-Schüler-Dialoge mit derselben Schülerin bzw. demselben Schüler. Dabei fällt auf, dass die („End“)Stationen 3, 5, 7, 9 und 11 (Beendigung durch die Lehrkraft) häufiger als die benachbarten Stationen vorkommen. Mit anderen Worten: Das letzte Wort haben bei Dialogen überwiegend die Lehrpersonen.

Diese Art von Analysen verdeutlicht das Potenzial videobasierter Analyse von Lehrer-Schüler-Interaktionen: die Identifikation von zeitlichen Mustern, Sequenzen und Skripts.

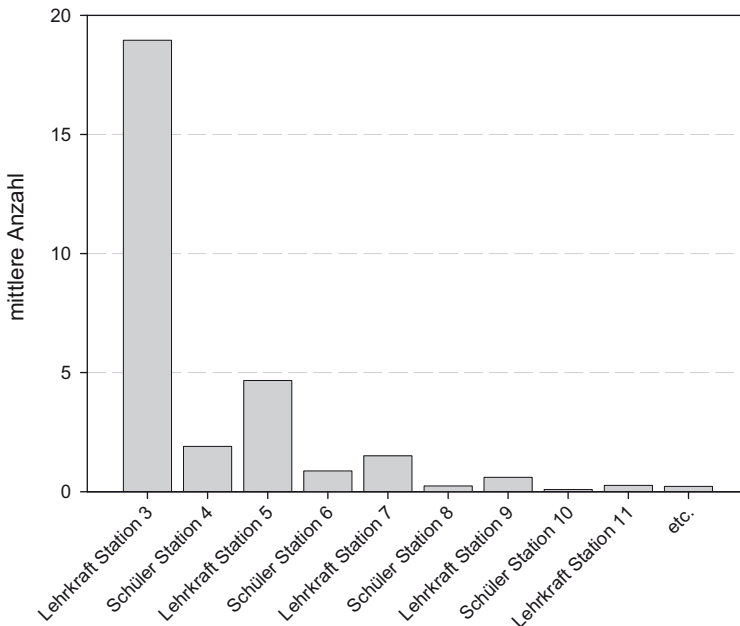


Abbildung 29.6: Häufigkeit der Lehrer-Interaktionen mit dem gleichen Schüler / der gleichen Schülerin im Englischunterricht.

29.4 Unterrichtswirksamkeit

Bislang ging es ausschließlich um die Beschreibung verschiedener Aspekte des Unterrichtsgeschehens. Welche Bedeutung haben aber nun bestimmte Unterrichtsmerkmale für den Lernerfolg der Schülerinnen und Schüler? Um dieser Frage nachzugehen, werden Zusammenhänge zwischen ausgewählten Unterrichtsmerkmalen und verschiedenen Zielkriterien berechnet. Die Zielkriterien werden dabei auf Klassenebene analysiert. Im Folgenden verwenden wir das *Hörverstehen* und die *Textrekonstruktionsfähigkeit* (C-Test), für die jeweils Zweipunktmessungen vorliegen, als Zielkriterien.

Vorhersage des Hörverstehens und der Textrekonstruktionsleistung

Mit den folgenden Merkmalen verbindet man in der allgemeinen Unterrichtsforschung und in der Fremdsprachendidaktik eine lernförderliche Wirkung:

- Langsamkeitstoleranz: Wartezeit auf Schülerantworten
- Hoher zeitlicher Anteil der Schülersprechzeit an der Gesamtredezeit
- Verwendung der englischen statt der deutschen Sprache im Unterricht
- Aufgabenorientierung und Strukturiertheit
- Schüleraktivierung: Einbezug möglichst aller Schüler in die Kommunikation
- Mehr längere Schüleräußerungen („sustained speech“), weniger „Ein-Wort-Sätze“

Tabelle 29.3 stellt die korrelativen Ergebnisse zur Rolle dieser Merkmale dar. Mithilfe des Programms Mplus wurden auf Klassenebene verschiedene Arten von Korrelationen zwischen Unterrichtsmerkmal und dem Nachtest (*Hörverstehen*) berechnet. Neben (a) der einfachen Korrelation berichten wir Partialkorrelationen⁵ (vgl. Kapitel 26) der Unterrichtsmerkmale (b) mit dem Nachtest und (c) dem Zuwachs.

Tabelle 29.3 zeigt, dass sich in den meisten Fällen signifikante positive oder negative Zusammenhänge in der erwarteten Richtung ergeben: Für Lernzuwachs sorgen häufige Sprechgelegenheiten für die Schüler, die Benutzung des Englischen, die Selbstkorrektur von Fehlern, ein didaktisch angemessenes Anforderungsniveau und das allgemeine Engagement der Schülerschaft. Negative Zusammenhänge zeigen sich bei der Verwendung der deutschen Sprache im Englischunterricht und bei Ein-Wort-Sätzen.

Es finden sich jedoch nicht vorausgesagte und erwartungswidrige Ergebnisse. Dazu gehört, dass die Verwendung von Deutsch als Unterrichtssprache außerhalb des Unterrichtsgesprächs (hier: in Übergangsphasen) nicht nur keinen negativen, sondern sogar einen positiven Effekt auf die Leistungsentwicklung hat (ohne Tabelle), und dass engführende Fragen – anders als im Mathematikunterricht – im Fremdsprachenunterricht sich leistungsförderlich auswirken.

Weiterhin war angenommen worden, dass sich die Strukturiertheit förderlich auf den Kompetenzerwerb auswirken würde. Die Ergebnisse zeigen jedoch, dass dies nicht der Fall ist. Das Ausmaß der Strukturiertheit des Englischunterrichts – egal ob mit der Basiskodierung oder mit dem Rating erhoben – korreliert weder mit dem Zuwachs noch mit dem Nachtest statistisch signifikant, ist also praktisch bedeutungslos.

Exemplarisch sei die Bedeutung der Korrelationen am Beispiel des „nicht-fachlichen Unterrichtsgegenstandes“ demonstriert:

- In leistungsstärkeren Klassen wird weniger Zeit in nicht-fachliche Aktivitäten investiert (einfache Korrelation: $r = -.28$);
- Je geringer der Zeitanteil an fachfremden Aktivitäten, desto höher ist der Leistungszuwachs beim Hörverstehen (Partialkorrelation mit dem Zuwachs: $r = -.42$);

5 Auspartialisiert wurden in den folgenden Tabellen (a) auf Individualebene: Leistungen zu Beginn der neunten Jahrgangsstufe, Geschlecht und Erstsprache sowie (b) auf Klassenebene: Leistungen zu Beginn der neunten Jahrgangsstufe, Geschlecht, Erstsprache und Bildungsgang (RS, GY). Die zusätzliche Berücksichtigung der kognitiven Grundfähigkeit als Kontrollvariable führte dazu, dass die Modelle nicht konvergierten, so dass hier nicht die sonst üblichen, nur die eben aufgeführten Kontrollvariablen verwendet wurden.

- Je geringer der Zeitanteil an fachfremden Aktivitäten, desto höher ist die um die Eingangsbedingungen und relevante Kontextbedingungen bereinigte Nachtestleistung (Partialkorrelation mit dem Nachtest: -.19).

Tabelle 29.3: Korrelationen videobasierter Unterrichtsmerkmale mit Hörverstehen, geordnet nach der Höhe (Betrag) der Partialkorrelation mit dem Zuwachs. Nicht signifikante Korrelationen ($p > .05$) sind nicht aufgeführt.

	Einfache Korrelation mit dem Nachtest	Partial- korrelation mit dem Zuwachs	Partial- korrelation mit dem Nachtest
Basiskodierung auf Mikroebene			
Nicht-fachlicher Unterrichtsgegenstand ^A	-.28*	-.42*	-.19*
Lehrer spricht deutsch im Unterrichtsgespräch		-.34***	-.15***
Sprechanteil für Schüler/in	.29**	.30**	.14**
Sprechen nach Vorgabe		.31*	.14*
Selbstkorrektur von Fehlern	.25*	.28**	.13**
Schüler/in spricht Englisch	.65***	.30*	.13*
Wartezeit nach Lehrerfragen		.27*	.12*
Ein-Wort-Sätze	-.45***	-.27**	-.12**
Ablesen eigener (Schüler-)Texte		.26*	.12*
Gleicher Schüler spricht mit Lehrer	.36**	.23*	.10*
Lehrer spricht deutsch in <i>Übergangsphase</i>		.17*	.08*
Ratings der Stunde			
Passung: Angemessenes Anforderungsniveau	.51***	.38**	.17**
Lenkung: Engführung von Fragen		.34*	.16*
Klassenführung: Aufgabenorientierung	.35**	.34*	.16*
Klassenführung: Zeitznutzung	.33*	.28*	.13*
Klassenführung: Störungsfreiheit	.39**	.28*	.13*
Motivierungsqualität: Schülerengagement	.56***	.23*	.11*
Fehlerklima: Positiver Umgang der Lehrperson mit Schülerfehlern	.34**	.21*	.10*

Anmerkung: ^ADass die (einfache und partielle) Korrelation dieses Merkmals mit dem Nachtest trotz numerisch höherer Werte, verglichen mit anderen Variablen, statistisch nicht oder nicht so klar gesichert werden kann, hängt mit der anormalen Verteilungsgestalt des Merkmals zusammen (ausgeprägte Schiefe und Exzess). Signifikanzniveaus: * $p < .05$; ** $p < .01$; *** $p < .001$.

Ein plastisches Beispiel dafür, was man sich unter „nicht-fachlichem Unterrichtsgegenstand“ vorstellen kann, liefert der folgende Auszug aus dem Transkript einer videografierten Englischstunde in Tabelle 29.4.

Tabelle 29.4: Beispiel für nicht-fachlichen Unterrichtsgegenstand (Transkription); S = Schüler, L = Lehrer.

21:43 - 21:46	S	Können wir das Fenster aufmachen, bitte?
21:46 - 21:47	S	Das stinkt hier.
21:47 - 21:48	L	Machen wir gleich in der Pause.
21:48 - 21:50	S	Wieso Pause, ich krieg Kopfschmerzen hier.
21:50 - 21:51	L	Ja, macht nix.
21:51 - 21:52	S	Macht nix, nein?
21:52 - 21:57	L	Nö, bringt doch Abwechslung, wenn mal einer umfällt, ist überhaupt kein Thema.

Die Zusammenhänge der videobasierten Unterrichtsmerkmale mit dem C-Test weisen ein ähnliches Muster auf. In Tabelle 29.5 werden lediglich diejenigen Unterrichts- und Lehrermerkmale aufgeführt, die statistisch signifikant mit dem bereinigten Nachtest korrelieren – und dies sind andere Merkmale als diejenigen, die für das Hörverstehen relevant sind. Auch hier wieder wirkt sich die Unterrichtssprache Deutsch im Englischunterricht durchgehend negativ aus, während die Klarheit des Unterrichts, ein positives Sozialklima oder auch das Anspruchsniveau positiv korrelieren.

Tabelle 29.5: Korrelationen videobasierter Unterrichtsmerkmale mit dem C-Test.

	einfache Korrelation mit dem Nachtest	Partial- korrelation mit dem Zuwachs	Partial- korrelation mit dem Nachtest
Lehrer spricht deutsch	-.53**	-.34*	-.08*
Einfachheit von Fragen	-.04	.30*	.07*
Klarheit	.51**	.32**	.07**
Anspruchsniveau	.49***	.31*	.07*
Sozialklima	.39**	.23*	.05*
Förderungsorientierung (Eingehen auf Schwierigkeiten)	.28*	.23*	.05*
Positives Feedback	.00	.21*	.05*
Lehrerengagement	.38**	.21*	.05*
Bekundung von Nicht-Wissen	.03	.17*	.04*

Anmerkung: Signifikanzniveaus: * $p < .05$; ** $p < .01$; *** $p < .001$.

Erwartungsgemäß gibt es erhebliche Unterschiede zwischen den einfachen und den kontextbereinigten Korrelationen. Beim Merkmal „Schüler sprechen Deutsch“ (siehe Tabelle 29.3) lässt sich dies besonders gut veranschaulichen: Die hohe *einfache* Korrelation von $r = .65$ zwischen der durchschnittlichen Englischkompetenz (Nachtest) und dem Prozentanteil „Schüler sprechen Englisch“ besagt lediglich, dass in leistungsstärkeren Klassen im Durchschnitt mehr Englisch gesprochen

wird. Dies ist trivial, weil es natürlich die Bildungsgangunterschiede widerspiegelt. Bereinigt man den Zusammenhang zwischen Unterrichtsmerkmal und Testleistung um den Bildungsgang und andere Kontextmerkmale und berücksichtigt man, dass die Unterrichtsqualität ihrerseits auch vom Kontext beeinflusst wird, dann ist der Zusammenhang zwischen Unterrichtsmerkmal und Testleistung zwar immer noch statistisch signifikant, mit $r = .13$ jedoch wesentlich schwächer (*partielle Korrelation*).

29.5 Komplexe Modelle

Bisher wurden lediglich bivariate Zusammenhänge einzelner Merkmale von Lehrpersonen bzw. des Unterrichts modelliert. Komplexe Modellierungen beziehen simultan mehrere Prädiktoren in die Vorhersage der Kriteriumsleistung ein und berücksichtigen dabei zwei Prämissen: Zum einen sollen komplexe Modelle erklärungsstark sein, zum anderen sparsam. Zur Gewährleistung einer starken Erklärung ist es nötig, sich auf Merkmale zu beschränken, deren bivariater Zusammenhang mit dem Zielkriterium bereits empirisch belegt wurde. Um die unerwünschten statistischen Folgen der Multikollinearität zu vermeiden, sollten für komplexe Analysen möglichst Unterrichtsmerkmale ausgewählt werden, die sich in vorangegangenen Studien als theoretisch und empirisch eigenständig erwiesen haben, hier z.B. der Sprechanteil der Schüler (als Merkmal der *Aktivierung*) und die Wartezeit nach Lehrerfragen (als Merkmal eines *lernförderlichen Klimas*). Wenn man diese beiden *fachübergreifenden* Konstrukte mit dem *fachspezifischen* Qualitätsmerkmal „Unterrichtssprache Englisch“ koppelt, resultiert das im Folgenden dargestellte Pfadmodell, das die Hörverstehensleistung im Englischen am Ende der neunten Jahrgangsstufe durch drei zentrale videobasierte Unterrichtsmerkmale als Prädiktoren und die üblichen Kontrollvariablen zu erklären versucht.

Abbildung 29.7 zeigt, dass es sich bei dem Schüler-Sprechanteil, der Verwendung von Englisch als Unterrichtssprache und der Wartezeit nach Lehrerfragen um drei voneinander weitgehend unabhängige Unterrichtsmerkmale handelt, die gleichermaßen einen Erklärungsbeitrag für das Hörverstehen am Ende der neunten Jahrgangsstufe liefern. Die Rangordnung der Hörverstehenskompetenz bleibt im Zeitraum zwischen Anfang und Ende der neunten Jahrgangsstufe weitgehend stabil, d.h. Schulklassen, die zu Beginn der neunten Jahrgangsstufe ein hohes oder niedriges Leistungsniveau aufweisen, zeichnen sich auch am Ende der neunten Jahrgangsstufe durch hohe oder niedrige Leistungen aus.

Die Leistung zu Beginn der neunten Jahrgangsstufe hat jedoch nicht nur einen Effekt auf die Leistung am Ende, sondern wirkt sich massiv auch auf den Unterricht aus. Während der Sprechanteil der Schülerinnen und Schüler nicht von deren Leistungsniveau im Hörverstehen zu Beginn der 9. Klasse abhängt, hat dieses einen hohen Einfluss auf die verwendete Unterrichtssprache und einen mittelhohen Einfluss auf die Wartezeit. Englisch wird als Unterrichtssprache eher eingesetzt, wenn die Klasse günstige Voraussetzungen im Hörverstehen aufweist. In diesen Klassen sind Lehrkräfte auch eher bereit, nach einer an Schülerinnen und Schüler gerichteten Frage auf die Antwort zu warten.

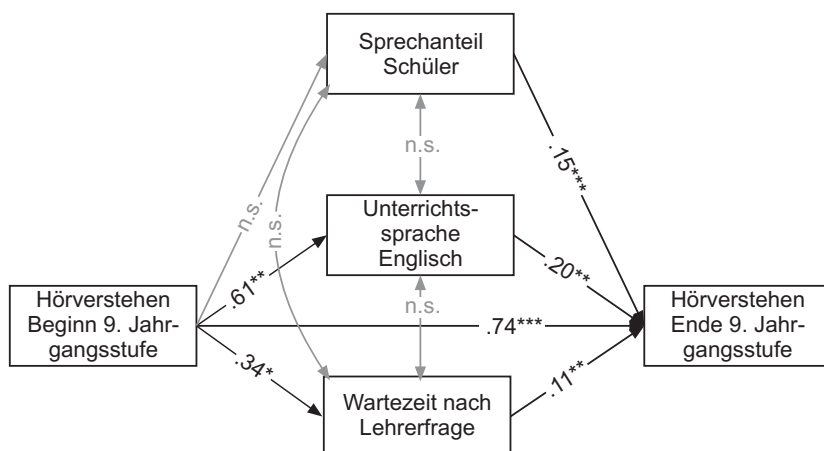


Abbildung 29.7: Vorhersage der Hörverstehensleistung am Ende der neunten Jahrgangsstufe durch die Eingangsleistungen zu Beginn der neunten Jahrgangsstufe und Merkmale der Unterrichtsqualität.

29.6 Wechselwirkungen

Im Zusammenhang mit der Heterogenität der kognitiven Eingangsvoraussetzungen ist ein weiteres Ergebnis von DESI von Bedeutung, und zwar zu der Frage: Profitieren eigentlich alle Schülerinnen und Schüler einer Klasse in gleicher Weise von gutem Unterricht bzw. leiden sie in gleicher Weise unter schlechtem Unterricht? In vielen Untersuchungen wie in pädagogischen Diskussionen ist dies eine unausgesprochene und meist nicht hinterfragte Prämisse.

In DESI sind wir dieser Frage exemplarisch für die kognitive Grundfähigkeit (gemessen mit dem Kognitiven Grundfähigkeitstest KFT) nachgegangen. Es ist ebenso bekannt wie trivial, dass Intelligenz und Schulleistung eng zusammenhängen. Wie aber wirkt sich die kognitive Grundfähigkeit auf den Leistungszuwachs aus, und wovon hängt die Enge dieses Zusammenhangs ab? Die DESI-Ergebnisse der Videostudie zeigen: Kognitive Grundfähigkeit und Leistungszuwachs (beim C-Test) hängen auf Individualebene im Durchschnitt schwach positiv zusammen. Die Höhe dieses Zusammenhangs variiert jedoch in Abhängigkeit von Merkmalen des Unterrichts und der Klassenzusammensetzung: Je unklarer und unverständlicher der Unterricht ist (Basis ist das Unterrichts-Rating in der Videostudie), desto stärker wird der Einfluss von Intelligenzunterschieden auf den Leistungszuwachs. Mit anderen Worten: Die unterschiedliche kognitive Grundfähigkeit der Schülerinnen und Schüler schlägt um so mehr durch, je höher das kognitive Niveau der Klasse und je geringer die Qualität des Unterrichts ist. Ist der Unterricht ausgesprochen klar, verständlich und schülerorientiert, dann kommt kognitiven Fähigkeitsunterschieden eine geringere Bedeutung für den Lernerfolg zu. Unterricht kann also je nach seiner Qualität dazu beitragen, den Zusammenhang zwischen kognitiver Grundfähigkeit

und Leistungsentwicklung entweder zu verschärfen oder zu entkoppeln. Die Vorstellung, dass sich der Unterricht auf Schülerinnen und Schüler mit heterogenen Eingangsvoraussetzungen in gleicher Weise auswirkt, kann unseren Ergebnissen zufolge für die Entwicklung der Sprachkompetenz im Englischen bei DESI nicht aufrechterhalten werden.

29.7 Zusammenfassung und Ausblick

Fasst man die videobasierten Analysen zusammen, dann lässt sich der Unterricht in Klassen mit besonders hohem Zuwachs in der *Hörverstehensleistung* wie folgt charakterisieren: In erfolgreichen Klassen

- kommen Schülerinnen und Schüler häufig zum Sprechen (Verteilung der Sprechanteile)
- warten Lehrerinnen und Lehrer mindestens drei Sekunden auf Schülerantworten
- ist bei Unterrichtsgesprächen Englisch die überwiegende Unterrichtssprache
- erhalten Schülerinnen und Schüler Gelegenheit zur Selbstkorrektur ihrer Fehler
- gibt es vergleichsweise wenige „Ein-Wort-Sätze“
- kommt es zu Lehrer-Schüler-Dialogen (mehrere Gesprächsstationen, über einfache Frage-Anwort-Sequenzen hinaus)

Daneben bestätigen sich einige Befunde, die bereits aus der fragebogenbasierten Hauptuntersuchung zum Unterricht berichtet wurden: Erfolgreicher Unterricht (bezogen auf Hörverstehens-Zuwachs) ist gekennzeichnet durch intensive Zeitznutzung und ausgeprägte Aufgabenorientierung, Störungsfreiheit, ein positives Fehlerklima und ein hohes Schülerengagement während des Unterrichts.

Interessant ist aber auch ein Blick auf diejenigen Unterrichtsmerkmale, die entweder keinen bedeutsamen Beitrag zur Erklärung von Leistungsunterschieden leisten oder deren Erklärungsbeitrag auf den ersten Blick womöglich überraschend ist:

- Keinen nachweislichen Effekt – weder auf die Englischleistung noch auf die Einstellung zum Fach Englisch – haben *Merkmale der Lehrerfrage* wie z.B. Antwortspielraum oder Authentizität. Anders als im Mathematikunterricht wirken sich enggeführte Fragen (mit geringem Antwortspielraum) durchaus positiv auf die Entwicklung des Hörverstehens aus. Dies dürfte damit zusammenhängen, dass Formulierungarbeit ein bedeutender Bestandteil von Englischunterricht ist.
- Im Gegensatz zu Untersuchungen des Mathematikunterrichts kommt dem Ausmaß der *Strukturierung* (durch Vorausschau, Zusammenfassungen etc.) keine positive Rolle zu. Dies wird verständlich, wenn berücksichtigt wird, dass der Erwerb primär von fremdsprachlich-kommunikativen Kompetenzen, nicht jedoch vom systematischen Aufbau von Wissensstrukturen abhängt, wie besonders auch Untersuchungen zum natürlichen Spracherwerbsprozess zeigen.
- Einen für das Hörverstehen leistungsförderlichen Effekt hat das *Vorlesen* eigener, d.h. selbst produzierter Texte.

Wir konnten in diesem Kapitel nur einen kurzen Einblick in Ergebnisse der Videostudie geben. Weitere Publikationen, die die Ergebnisse dieser Studie nutzen, werden sich mit folgenden Themen beschäftigen:

- Analyse von Zusammenhängen zwischen fragebogenbasierten (Schülerangaben) und videobasierten Angaben zur Unterrichtsqualität.
- Integrative Analysen von Daten, die aus mehreren Perspektiven (Lehrer vs. Schüler) erhoben wurden.
- Einbezug von Ergebnissen der Schülerbefragung unmittelbar nach der videograferten Unterrichtsstunde als zusätzlicher Erklärungsvariable (Mediationsprozesse).
- Identifikation von weiteren zeitlichen Abfolgen: Sequenzen, Muster, Skripte im Verlauf von Englischlektionen.
- Zusätzlich zu den hier berücksichtigten Zielkriterien: Zugrundelegung weiterer Indikatoren der Englischleistung (z.B. SET 10-Test) und zusätzlicher nicht-kognitiver Schülermerkmale (z.B. Realitätsangemessenheit der Selbsteinschätzung).
- Nonverbale und paraverbale Kommunikation im Englischunterricht.
- Classroom discourse analysis.
- Ausschöpfung der gesamten Breite der Ratings (z.B. auch Humor).
- Vertiefende Kodierung der Qualität (inhaltlich und sprachlich) der Lehreräußerungen durch *native speakers*.
- *Best Practice*: Ermittlung des Unterrichtsprofils von Klassen mit erwartungswidrig gutem Leistungsergebnis.

Literatur

- Baumert, J./Lehmann, R. H./Lehrke, M./Schmitz, B./Clausen, M./Hosenfeld, I./Köller, O./Neubrand, J. (Hrsg.) (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich.
- Bausch, K. R./Christ, H./Krumm, H. J. (Hrsg.) (2002): Handbuch Fremdsprachenunterricht. Tübingen: Francke.
- Helmke, A. (2004): Unterrichtsqualität: Erfassen, Bewerten, Verbessern. Seelze: Klett-Kallmeyer.
- Helmke, A./Schrader, F.-W. (1993): Was macht erfolgreichen Unterricht aus? Ergebnisse der Münchner Studie. Praxis Schule, 5-10, H. 1, S. 11-13.
- Helmke, A./Schrader, F.-W. (1997): Unterrichtsbeurteilungen durch externe Beobachter. In: Weinert, F. E./Helmke, A. (Hrsg.): Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union, S. 510-518.
- Heuer, H./Klippel, F. (1993): Englischmethodik. Problemfelder, Unterrichtswirklichkeit, Handlungsempfehlungen. Berlin: Cornelsen.
- Hiebert, J./Gallimore, R./Garnier, H./Bogard Givvin, K./Hollingsworth, S./Jacobs, J./Chui, A. M. Y./Wearne, D./Smith, M./Kersting, N./Manaster, A./Tseng, E./Etterbeek, W./Manaster, C./Gonzales, P. & Stigler, J. W. (2003). Teaching Mathematics in Seven Countries: Results from the TIMSS 1999 Video Study. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Judd, E. L./Tan, L./Walberg, H. J. (2001): Teaching additional languages (Educational Practices Series, Vol. 6). Brussels: International Academy of Education & International Bureau of Education. URL: <http://www.ibe.unesco.org>.

- Klieme, E./Reusser, K. (2003): Unterrichtsqualität und mathematisches Verständnis im internationalen Vergleich – Ein Forschungsprojekt und erste Schritte zur Realisierung. In: *Unterrichtswissenschaft* 31, H. 3, S. 194-205.
- Klieme, E./Schümer, G./Knoll, S. (2001): Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In: Bundesministerium für Bildung und Forschung (Hrsg.): *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*. Bonn: Bundesministerium für Bildung und Forschung (BMBF), S. 43-58.
- Krashen, S. D./Terrell, T. (1983): *The Natural Approach. Language Acquisition in the Classroom*. Oxford: Pergamon, S. 59-61.
- Prenzel, M./Seidel, T./Lehrke, M./Rimmele, R./Duit, R. (2002): Lehr-Lernprozesse im Physikunterricht – eine Videostudie. In: Prenzel, M./Doll, J. (Hrsg.): *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen*. In: *Zeitschrift für Pädagogik* 45, Beiheft. Weinheim: Beltz, S. 139-156.
- Reusser, K./Pauli, C. (2004): Bericht über die Ergebnisse einer internationalen und schweizerischen Video-Unterrichtsstudie zum Mathematikunterricht in der Schweiz und in weiteren sechs Ländern. In: *SEMINAR* 10, H. 4, S. 102-104.
- Rimmele, R. (2005): *Videograph – Multimedia-Player zur Kodierung von Videos*. Kiel: Leibniz-Institut für die Pädagogik der Naturwissenschaften.
- Spada, N./Fröhlich, M. (1995): *The communicative orientation of language teaching (COLT) observation scheme: coding conventions and applications*. Sydney: National Centre for English Language Teaching and Research.
- Stigler, J. W./Gallimore, R./Hiebert, J. (2000): Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS Video studies. In: *Educational Psychologist* 35, S. 87-100.
- Timm, J.-P. (Hrsg.) (1998): *Englisch lernen und lehren*. Berlin: Cornelsen.
- Waxman, H. C./Hilberg, R. S./Tharp, R. G. (2004): Future Directions for Classroom Observation Research. In: Hilberg, R. S./Waxman, H. C./Tharp, R. G. (Hrsg.): *Observational Research in U.S. Classrooms: New Approaches for Understanding Cultural and Linguistic Diversity*. Cambridge: Cambridge University Press, S. 266-277.
- Wellenreuther, M. (2004): *Lehren und Lernen – aber wie?* Baltmannsweiler: Schneider Verlag Hohengehren.
- Wirtz, M./Caspar, F. (2002): *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.

Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Günter Nold / Konrad Schröder

30 Soziodemografische und fachdidaktisch relevante Merkmale von Englischlehrpersonen

In diesem Kapitel geht es – anknüpfend an das im Kapitel 26 dargestellte Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts – um Merkmale von Lehrpersonen, die als Bedingungen unterrichtlichen Handelns gelten können und daher in der Wirkungskette noch vor der Qualität und Quantität des Unterrichts anzusiedeln sind. Neben ausgewählten soziodemografischen und berufsbiographischen Merkmalen sind dies Ausbildung, Fortbildung, Aspekte der fachdidaktischen, sprachlichen und pädagogischen Expertise sowie die Orientierung in Richtung englischsprachiges Ausland. Aus Platzgründen müssen wir uns auf eine Auswahl der bei DESI insgesamt erfassten Merkmale von Lehrpersonen beschränken. Es geht im Folgenden ausschließlich um die *Beschreibung* der genannten Personmerkmale. *Zusammenhänge* zwischen den Lehrermerkmalen und dem Leistungsniveau der Schülerinnen und Schüler sind dann Gegenstand des Kapitels 32 zur Wirksamkeit von Person- und Unterrichtsmerkmalen.

30.1 Stichprobenbeschreibung

An der Befragung nahmen zu Beginn der neunten Jahrgangsstufe 366 und am Ende 328 Englisch-Lehrkräfte teil¹. Legt man die Beteiligung am Ende der neunten Jahrgangsstufe zugrunde, dann lässt sich die Zusammensetzung der Stichprobe (in Prozent der befragten Lehrpersonen) wie folgt beschreiben:

- *Schulart*: Hauptschule (53 Klassen, 16%), Realschule (89 Klassen, 27%), Gymnasium (137 Klassen, 42%), Integrierte Gesamtschule (15 Klassen, 5%) sowie Schulen mit mehreren Bildungsgängen (34 Klassen, 10%, davon 10 im Bildungsgang Hauptschule und 24 im mittleren Bildungsgang).
- *Geschlecht*: 68% Frauen, 32% Männer.
- *Alter*: bis 31 J. (7%), 31-35 J. (12%), 36-40 J. (11%), 41-45 J. (13%), 46-50 J. (17%), 51-55 J. (26%), 56-60 J. (10%) und über 60 J. (6%).
- *Beschäftigungsstatus*: Vollzeit (68%), Teilzeit (32%).

Für alle inhaltlichen Aussagen über Merkmale von Lehrpersonen ist zweierlei zu beachten: (1) Bei den im Folgenden berichteten Ergebnissen handelt es sich stets um *gewichtete Daten*, so dass die getroffenen Aussagen auf Englischunterricht in der

¹ Die Angaben beziehen sich auf die in DESI erhobenen 388 Klassen, bei denen alle Schülerinnen und Schüler der jeweiligen Klasse von derselben Englischlehrkraft unterrichtet werden. Nur diese Klassen werden bei den Analysen berücksichtigt.

neunten Jahrgangsstufe an deutschen Schulen verallgemeinerbar sind. (2) Aufgrund der hier zugrunde liegenden Gewichtungszurordnung auf Schülerebene handelt es sich bei den Prozentangaben genau genommen nicht um Aussagen über den Prozentanteil der Lehrpersonen, sondern um Aussagen über den Prozentanteil an Schülern, die von entsprechenden Lehrpersonen in den unterschiedlichen Bildungsgängen unterrichtet werden.

30.2 Ausbildung

Die Frage nach einer professionalisierenden Ausbildung der fremdsprachlichen Lehrerschaft ist seit langem ein strittiges Thema. Der Bereich ist in zahlreichen Publikationen, vor allem aus den 70er Jahren des 20. Jahrhunderts (Mainusch u.a. 1976; Schröder/Walter 1973, 1979) und aus jüngster Zeit (Schröder 1992, 2000) umfassend dokumentiert. Welche Rolle spielt die Art der Ausbildung für die Tätigkeit als Englischlehrkraft? Hier wurde zweistufig differenziert (vgl. Tabelle 30.1):

- (a) ob dieses Fach studiert wurde oder nicht,
- (b) falls ja: ob Englisch als Haupt- oder Neben- bzw. Erweiterungsfach studiert wurde; falls nein: welche Art der Qualifikation dann zugrunde liegt.

Tabelle 30.1: Ausbildung nach Bildungsgang (Angaben in Prozent).

Bildungsgang	Englisch-Studium			
	ja		nein	
HS	71.6		28.4	
RS	90.0		10.0	
IGS	100.0		0.0	
GY	98.0		2.0	
	Hauptfach	Nebenfach ^A	Nachqualifikation	Schnellkurs, Lehrgang, Umschulung
HS	77.8	22.1	19.5	80.5
RS	89.0	11.0	53.4	46.6
IGS	81.3	18.7	-	-
GY	93.6	6.4	91.0	8.9

Anmerkungen: ^Aoder Erweiterungsfach, Umschulungsfach.

Bildungsgangunterschiede bezüglich der Frage, ob Englisch überhaupt studiert wurde, zeigen sich erwartungsgemäß vor allem zwischen Gymnasium (Englischstudium bei 98%) und Hauptschule (knapp 72%)².

2 Chi²-Test für Bildungsgangunterschiede: $p < .001$; Hauptschule vs. Realschule bzw. Gymnasium: jeweils $p < .001$; Realschule vs. Gymnasium: $p < .01$.

30.3 Fortbildung

Wie häufig bilden sich Englischlehrkräfte fort? Die Frage bezog sich nicht auf irgendwelche Fortbildungen, sondern nur auf solche, die unmittelbar mit dem Englischunterricht zu tun haben, wie z.B. Grundsatzfragen zur Fachdidaktik oder Fragestellungen zur Unterrichtsmethodik. Die Ergebnisse zeigen, dass die Mehrheit der Englischlehrkräfte im laufenden Schuljahr an einschlägigen Fortbildungen teilgenommen hat (HS: 42%, RS: 60%, IGS: 75% und GY: 60%). Dies kann als Beleg dafür angesehen werden, dass die Mehrheit der Englischlehrkräfte den Erwartungen an eine regelmäßige Weiterbildung zu entsprechen versucht.

30.4 Orientierung am englischsprachigen Ausland

Bei der Professionalisierung von Fremdsprachenlehrkräften nimmt die Orientierung am englischsprachigen Ausland eine zentrale Stellung ein. Es gibt dazu eine Reihe empirischer Untersuchungen (vgl. z.B. Ehrenreich 2004). Im Rahmen des DESI-Projektes wurden folgende Aspekte mittels des Lehrerfragebogens untersucht:

- Häufigkeit und Dauer eigener Reisen ins englischsprachige Ausland,
- Kontakte ins englischsprachige Ausland: persönlich, telefonisch, brieflich, per Email; Einschätzung der Relevanz von Kontakten für den Unterricht.

Die überwiegende Mehrheit aller Lehrpersonen hält Kontakte ins englischsprachige Ausland für wichtig (55%) oder sogar für unabdingbar (37%). Da es in Deutschland bisher keine repräsentativen empirischen Untersuchungen zum Auslandsaufenthalt und zur Orientierung von Englischlehrpersonen am englischsprachigen Ausland gibt, berichten wir die diesbezüglichen Ergebnisse im Detail (vgl. Abbildung 30.1)³.

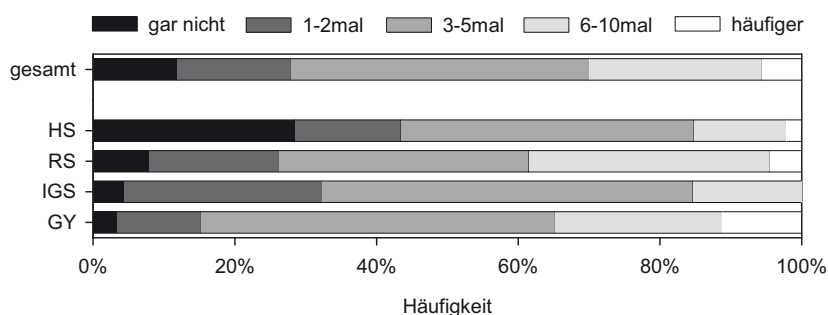


Abbildung 30.1: Häufigkeit von Aufenthalten im englischsprachigen Ausland innerhalb der letzten 10 Jahre.

Sie machen deutlich, dass insgesamt gesehen eine große Mehrheit der Englischlehrkräfte (88%) im letzten Jahrzehnt zumindestens 1-2 mal im englischsprachigen Ausland gewesen ist, einige sogar wesentlich häufiger. Bildungsgangunterschiede (mittels Varianzanalyse festgestellt, $p < .01$) finden sich insbesondere zwischen Hauptschule und Gymnasium ($p < .001$). Als Dauer des längsten Auslandsaufenthaltes

3 Aus Gründen der Übersichtlichkeit wurden die offen erfassten Antworten zusammengefasst.

gibt die überwiegende Mehrheit (82%) der Lehrkräfte mit mehr als zehn Jahren Berufserfahrung einen Zeitraum von bis zu vier Wochen an. Bemerkenswerterweise gibt es jedoch auch Lehrkräfte, die sich 1-3 Monate (7%), 4-6 Monate (3%) oder noch länger (8%) im englischsprachigen Ausland aufgehalten haben.⁴

Die Abbildungen 30.2 und 30.3 beschreiben die Häufigkeit verschiedener Modi des persönlichen Kontaktes mit dem englischsprachigen Ausland, insgesamt und separat für die Bildungsgänge.

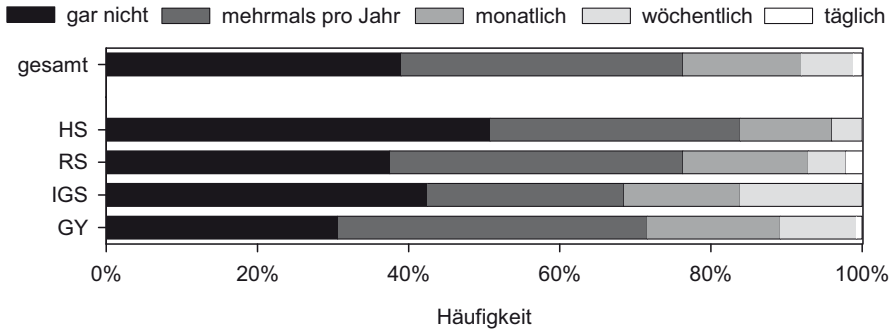


Abbildung 30.2: Häufigkeit von E-Mail-Kontakten ins englischsprachige Ausland.

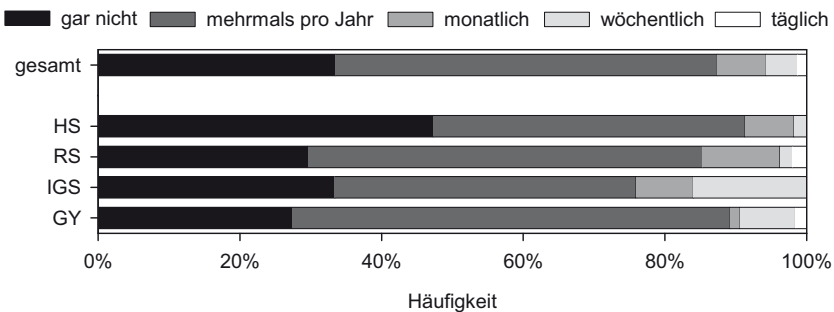


Abbildung 30.3: Häufigkeit persönlicher („face-to-face“) Kontakte mit Personen aus dem englischsprachigen Ausland.

Auch hier bestätigt sich das Bild, dass sich eine deutliche Mehrheit der Englischlehrkräfte aktiv um eine Orientierung an der Sprache des unterrichteten Faches bemüht. Die deutlichsten Kontraste zeigen sich zwischen den Bildungsgängen Hauptschule und Gymnasium; letztere pflegen häufigere Kontakte ins englischsprachige Ausland (Email- bzw. persönliche Kontakte: $p < .01$; telefonische Kontakte: $p < .001$).

4 In vielen Bundesländern wird ein Mindestaufenthalt im englischsprachigen Ausland als obligatorischer Teil der zweiten Phase der Lehrerausbildung angesehen. Die Analysen zu Zusammenhängen zwischen Auslandsaufenthalten und Unterrichts- sowie Leistungsmerkmalen der Klasse in Kapitel 32 berücksichtigen deshalb nur diejenigen Lehrkräfte, die seit mindestens zehn Jahren im Lehramt berufstätig sind.

30.5 Fachliches und pädagogisches Engagement

In der Pädagogischen Psychologie und Unterrichtsforschung geht man davon aus, dass die *Expertise* von Lehrkräften, d.h. die Menge, Qualität, Aktualität und Vernetztheit ihres unterrichtsrelevanten Wissens, eine wichtige Voraussetzung für das Gelingen des Unterrichts ist (Bromme 1997; Krauss u.a. 2004). Der Aufbau solcher für das praktische Handeln nutzbarer Wissenssysteme hängt entscheidend von der Art und dem Umfang der Erfahrungen ab, die im Zuge der Auseinandersetzung mit beruflichen Aufgabenstellungen erworben werden. Insofern sind aus unserer Sicht sowohl die über das Unterrichten hinausgehende Beschäftigung mit unterrichtsrelevanten Themen als auch spezielle Funktionen im Rahmen des Kollegiums wichtige Quellen der Expertise. Bei DESI wurden verschiedene Grundlagen und Indikatoren der Expertise mit folgenden Fragen erfasst:

Fachdidaktisches Engagement. Hier fragten wir danach, ob eine fachdidaktische oder auf Fragen des Unterrichts bezogene Zeitschrift regelmäßig bezogen wird. Dies bejaht insgesamt weniger als die Hälfte der befragten Lehrkräfte (HS: 26%, RS: 37%, IGS: 62%, GY: 53%)⁵.

Orientierung an der englischen Sprache. Hier konnte aus drei Items eine kurze Skala gebildet werden, die akzeptable psychometrische Gütekriterien aufweist (Cronbach's $\alpha = .70$) Häufigkeit der Lektüre englischsprachiger Zeitungen und Zeitschriften, des Anschauens englischsprachiger Sendungen im TV/Video und der Lektüre englischsprachiger Bücher. Die Orientierung an der englischen Sprache ist im Gymnasium erwartungsgemäß am höchsten ($M = 2.74$) und am niedrigsten ($M = 1.94$) in der Hauptschule. Die Abbildungen 30.4 bis 30.6 berichten die Ergebnisse der einzelnen Fragen im Detail, aufgeschlüsselt nach Bildungsgang.

Die Ergebnisse unterstreichen, dass die deutliche Mehrheit der Englischlehrkräfte zumindest gelegentlich (d.h. ein paar Mal pro Jahr) englischsprachige Zeitungen oder Zeitschriften liest, und dass noch mehr Lehrkräfte (80%) zumindest gelegentlich englischsprachige Filme oder Videos anschauen. Auch hier treffen wir auf deutliche Bildungsgangunterschiede⁶: Bei den englischsprachigen Zeitschriften unterscheiden sich alle Bildungsgänge – Ausnahmen stellen lediglich die Kontraste zwischen Hauptschule bzw. Realschule und Integrierte Gesamtschule dar – jeweils signifikant voneinander (jeweils $p < .001$). Vor allem im Gymnasium ist es überaus selten, dass Lehrkräfte überhaupt keine englischsprachigen Zeitungen oder Zeitschriften lesen.

5 Die Bildungsgangunterschiede sind hoch signifikant ($p < .001$). Bei den einzelnen Kontrasten gibt es lediglich keinen statistisch signifikanten Unterschied zwischen Hauptschule und Realschule bzw. Integrierter Gesamtschule und Gymnasium.

6 Chi²-Tests auf der Basis der dichotomisierten (,nie' bzw. ,nie oder fast nie' vs. restliche Kategorien) Merkmale, jeweils $p < .001$.

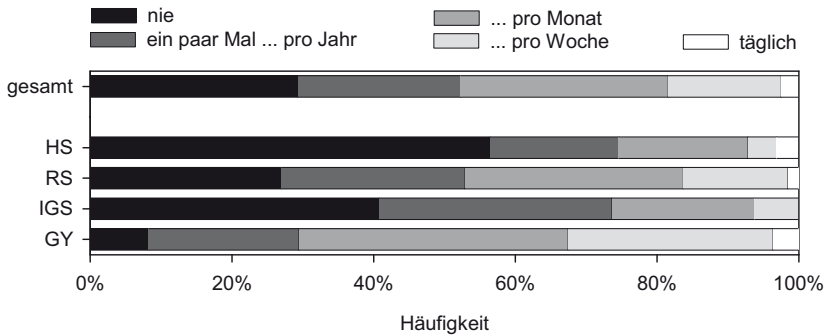


Abbildung 30.4: Häufigkeit der Lektüre englischsprachiger Zeitungen und Zeitschriften.

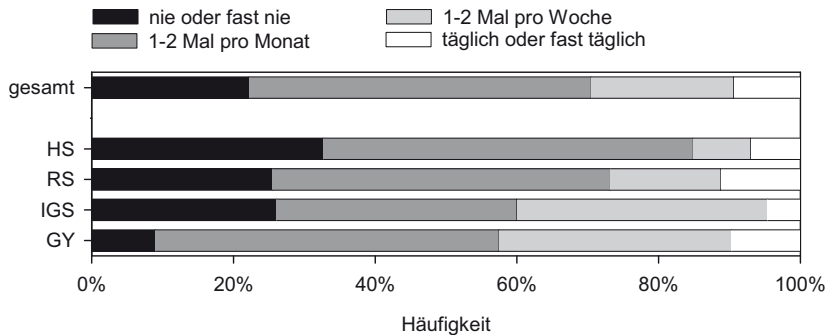


Abbildung 30.5: Häufigkeit des Verfolgens englischsprachiger TV-Sendungen oder Videos.

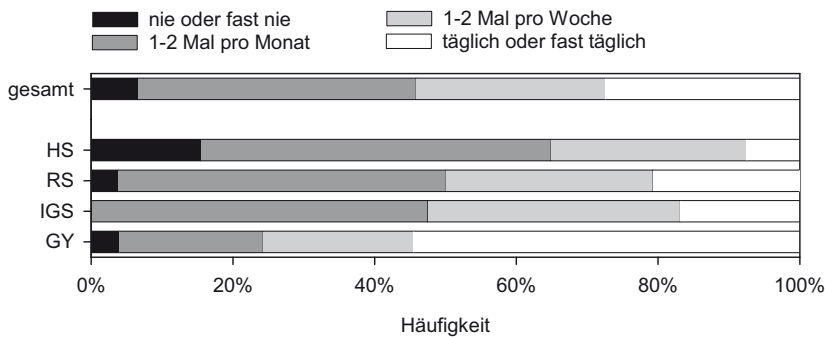


Abbildung 30.6: Lektüre englischsprachiger Texte.

Pädagogisches Engagement: Wir gehen davon aus, dass die Übernahme von Funktionen in der Ausbildung von Referendaren, Praktikanten oder Lehramtsanwärtern (die Bezeichnungen hierfür variieren von Bundesland zu Bundesland und von Schulart zu Schulart) als Mentor oder Seminarleiter ein Beleg für die pädagogische Kompetenz ist. Insgesamt gibt ein knappes Drittel der befragten Lehrkräfte (gewichtet: 29%) an, dass sie eine solche Funktion ausüben.

30.6 Ausblick

Insgesamt gesehen zeichnen die berichteten Ergebnisse ein durchaus positives Bild des fachlichen und fachdidaktischen Engagements von Englischlehrpersonen, insbesondere was die Orientierung am englischsprachigen Ausland anbelangt. Es ist allerdings nicht zu übersehen, dass sich Schülerinnen und Schüler in Hauptschulen in einer vergleichsweise ungünstigen Lage befinden, was ihre Englischlehrkräfte betrifft: Diese erteilen den Englischunterricht teilweise ohne vorangegangenes Fachstudium und weisen eine deutlich größere Distanz zum englischsprachigen Ausland auf als Lehrkräfte in anderen Bildungsgängen. Künftige vertiefende Analysen werden zum einen Zusammenhänge zwischen Auslandsaufenthalt, Orientierung an der englischen Sprache und fachdidaktischem Engagement analysieren und zum anderen die Ergebnisse der Videostudie einbeziehen.

Literatur

- Bromme, R. (1997): Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In: Weinert, F. E. (Hrsg.): Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Pädagogische Psychologie, Band 3). Göttingen: Hogrefe, S. 177-212.
- Ehrenreich, S. (2004): Auslandsaufenthalt und Fremdsprachenlehrerbildung. München: Langenscheidt.
- Krauss, S./Kunter, M./Brunner, M./Baumert, J./Blum, W./Neubrand, M./Jordan, A./Löwen, K. (2004): COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In: Doll, J./Prenzel, M. (Hrsg.): Die Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung. Münster: Waxmann, S. 31-53.
- Mainusch, H./Mertner, E./Schmidt, S. J./Schröder, K. (Hrsg.) (1976): Lehrerfortbildung und Lehrerweiterbildung in der Bundesrepublik Deutschland: Modell Anglistik. Bern und Frankfurt a.M.: Lang.
- Schröder, K. (1992): Die Geschichte des Fremdsprachenunterrichts der DDR ist noch nicht geschrieben. Historiographische Reflexionen und der Versuch, ein Stück frühe Entwicklung des Fremdsprachenunterrichts in der DDR nachzuzeichnen. In: Schröder, K. (Hrsg.): Situation und Probleme des Fremdsprachenunterrichts und der Fremdsprachenlehrausbildung in den Neuen Bundesländern. Anspruch - Wirklichkeit - Entwicklungsmöglichkeiten. Kolloquium zu Kahla (Thüringen) vom 12. bis 14. Dezember 1991. Augsburg: Universität, S. 125-165.
- Schröder, K. (2000): Unmaßgebliche Anmerkungen zur Lehrerbildung in den modernen Fremdsprachen. In: Helbig, B./Kleppin, K./Königs, F. G. (Hrsg.): Sprachlehrforschung im Wandel. Beiträge zur Erforschung des Lehrens und Lernens von Fremdsprachen. Festschrift für Karl-Richard Bausch zum 60. Geburtstag. Tübingen: Stauffenberg, S. 87-96.
- Schröder, K./Walter, G. (1973): Fremdsprachendidaktisches Studium in der Universität. München: Strumberger.
- Schröder, K./Walter, G. (Hrsg.) (1979): Fachdidaktisches Studium in der Lehrerbildung: Englisch. München: Oldenbourg.

Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Günter Nold / Konrad Schröder

31 Alltagspraxis des Englischunterrichts

Wie sieht die Alltagspraxis des Englischunterrichts in deutschen Schulen aus? Da es bisher kaum empirische Evidenz zur Alltagspraxis des Englischunterrichts basierend auf quantitativen Untersuchungsmethoden gibt, berichten wir im Folgenden Ergebnisse zu einem breiten inhaltlichen Spektrum und legen dabei überwiegend Angaben von Lehrkräften zugrunde. Obwohl gelegentlich auch *Skalen* gebildet wurden, liegt der Schwerpunkt dieses Abschnitts auf der Charakterisierung des Englischunterrichts auf der Ebene einzelner *Items* („thick description“). Dagegen werden in den Kapiteln, in denen es um die Erklärung von Unterschieden im Leistungsniveau zwischen Klassen und Schülern geht, überwiegend Skalen zugrunde gelegt, die auf Schülerangaben basieren.

Um voreilige Aussagen zur Leistungsrelevanz verschiedener Aspekte der Unterrichtsgestaltung zu vermeiden, berichten wir in diesem Abschnitt noch keine Zusammenhänge mit der Englischleistung. Dies erfolgt in den Kapiteln zur Unterrichtswirksamkeit (Kapitel 32) und zur Videostudie (Kapitel 29).

Organisationsformen des Unterrichts

Welche zum Frontalunterricht und Unterrichtsgespräch alternativen Organisationsformen lassen sich im Englischunterricht der Sekundarstufe finden? Aus Abbildung 31.1 ist eine Reihe solcher Methoden und Lehr-Lern-Szenarien sowie ihre auf den Lehrerangaben basierenden Verwendungshäufigkeiten ersichtlich; die in DESI verwendeten Items wurden vom Projekt MARKUS (Mathematik-Gesamterhebung Rheinland-Pfalz; Helmke/Jäger 2002) übernommen.

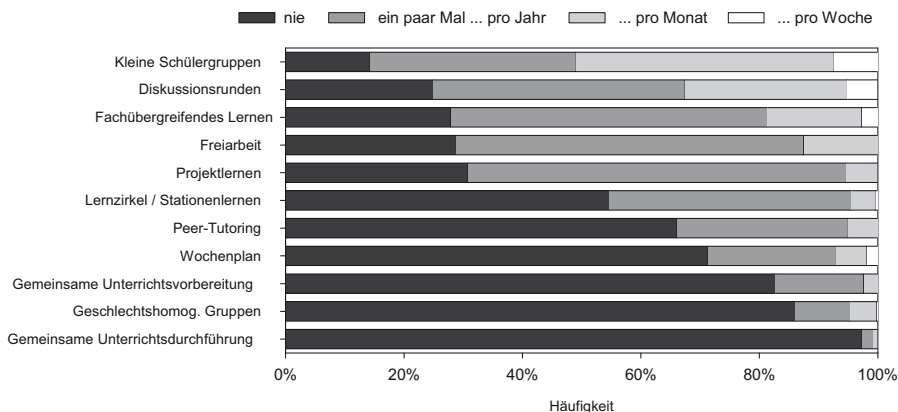


Abbildung 31.1: Häufigkeiten von Organisationsformen des Englischunterrichts jenseits des Frontalunterrichts (Lehrerangaben).

Die Ergebnisse unterstreichen, dass Kleingruppenarbeit die am häufigsten praktizierte Alternative zum Frontalunterricht darstellt, während Lehr-Lern-Szenarien wie Freiarbeit, Stationenlernen und Projektlernen schon deutlich seltener vorkommen. Von Pädagogen immer wieder gefordert, aber noch seltener als im Mathematikunterricht (Helmke/Hogenfeld/Schrader 2002) sind kooperative Formen der Unterrichtsvorbereitung oder gar -realisierung.

Tabelle 31.1 berichtet separat für die vier Bildungsgänge, wieviel Prozent der Lehrkräfte die jeweilige Organisationsform mindestens ein paar Mal pro Monat realisieren. Es zeigt sich, dass in der Hauptschule signifikant weniger Gruppenarbeit realisiert wird als in den anderen Bildungsgängen, während Wochenplanarbeit in der Integrierten Gesamtschule besonders häufig vorkommt. Wenn der Englischunterricht überhaupt gemeinsam vorbereitet wird, dann am ehesten noch in der Realschule.

Tabelle 31.1: Organisationsformen des Englischunterrichts nach Bildungsgang (in Prozent).

Organisationsform	HS	RS	IGS	GY	Gesamt
Arbeit mit kleinen Schülergruppen	31.9	52.6	69.9	61.5	50.9
Diskussionsrunden	4.1	40.2	39.4	45.8	32.5
Fachübergreifendes Lernen	13.0	21.9	25.8	17.4	18.6
Freiarbeit	17.2	7.8	18.9	13.7	12.5
Wochenplan	7.2	5.4	28.0	4.5	7.0
Projektlernen	1.6	7.5	0.0	6.7	5.3
Peer-Tutoring	4.9	3.4	4.9	7.1	5.0
Geschlechtshomogene Kleingruppen	1.5	4.6	10.6	6.1	4.6
Lernzirkel/Stationenlernen	9.7	2.3	9.2	1.6	4.4
Gemeinsame Unterrichtsvorbereitung	3.2	3.8	0.0	0.2	2.4
Gemeinsame Unterrichtsdurchführung	0.0	1.8	0.0	0.0	0.7

Anmerkungen: HS: Hauptschule; RS: Realschule; IGS: Integrierte Gesamtschule; GY: Gymnasium.

Neben dieser Detaildarstellung lässt sich auch – über alle hier aufgeführten Methoden hinweg – ein Gesamtwert bilden, den wir „*Methodenvielfalt*“ nennen. Dabei ergeben sich die folgenden Mittelwerte: 1.50 (HS), 1.65 (RS), 1.70 (IGS) und 1.65 (GY). In der Hauptschule ist die Vielfalt „alternativer“ Lehr-Lern-Formen im Englischunterricht demnach geringer als in den anderen Bildungsgängen – umgekehrt verglichen mit dem im Projekt MARKUS für Mathematik gefundenen Ergebnismuster, bei dem die Methodenvielfalt im Gymnasium mit Abstand am geringsten und in der Hauptschule am höchsten war (Helmke/Hosenfeld/Schrader 2002).

Der Informationswert dieser Skala soll jedoch aus zwei Gründen nicht überschätzt werden: Zum einen wäre es naiv anzunehmen, ein Unterricht sei um so besser, je mehr Methoden eingesetzt werden (Helmke 2004). Aus diesem Grunde rangieren die Ergebnisse zu den Organisationsformen auch unter „Alltagspraxis“ und nicht unter „Unterrichtsqualität“: „Methodenvielfalt“ ist nicht per se ein Gütemerkmal, weil es nicht auf die Quantität, sondern die Qualität, d.h. auf die gute Orchestrierung, den

angemessenen Mix, das gute Timing, die richtige Dosierung verschiedener Methoden ankommt.

Leistungsbeurteilung

Wie häufig werden im Englischunterricht unterschiedliche Formen von Lernerfolgskontrollen eingesetzt? Die Ergebnisse in Tabelle 31.2 berichten die absolute Anzahl im laufenden Schuljahr.

Tabelle 31.2: Lernerfolgskontrollen im Englischunterricht nach Bildungsgang (mittlere Häufigkeit) – geordnet in absteigender Reihenfolge.

	HS	RS	IGS	GY	Gesamt
Abfragen von Vokabeln	8.7	9.4	8.4	8.4	8.9
Mündliche Lernerfolgskontrollen	3.9	7.9	8.4	8.1	6.9
Klassenarbeiten	5.4	4.6	6.1	4.8	5.0
Informelle Tests (offene Antworten)	0.9	2.1	1.8	1.7	1.7
Vergleichsarbeiten	2.4	1.1	0.3	0.4	1.2
Informelle Tests (Multiple Choice)	1.0	1.4	1.1	1.0	1.2
Standardisierte Englischtests	0.6	0.7	0.6	0.5	0.6
Portfolio	0.3	0.4	0.6	0.2	0.3

Anmerkungen: HS: Hauptschule; RS: Realschule; IGS: Integrierte Gesamtschule; GY: Gymnasium.

Traditionelle Varianten der Leistungsüberprüfung rangieren demzufolge ganz vorn, während Vergleichsarbeiten und standardisierte Englischtests wenn überhaupt nur einmal pro Jahr eingesetzt werden. In der Hauptschule werden mündliche Lernerfolgskontrollen signifikant seltener, Vergleichsarbeiten signifikant häufiger als in anderen Bildungsgängen eingesetzt. Angesichts der aus den Bildungsstandards ableitbaren Aufforderung zur regelmäßigen Kompetenzdiagnose sowie der Lernstandserhebungen in den meisten Bundesländern ist in diesem Bereich in nächster Zeit mit gravierenden Änderungen zu rechnen.

Differenzierung

Seit PISA steht die Frage nach dem Umgang mit Leistungsheterogenität im Zentrum vieler Diskussionen und Recherchen. Bei DESI wurde nach der Häufigkeit gefragt, mit der verschiedene Formen der inneren Differenzierung im Unterricht realisiert werden. Obwohl auch hier ein Gesamtwert (Häufigkeit unterschiedlicher Differenzierungsarten) gebildet wurde, sind gerade die Informationen über die Art und Weise, wie differenziert wird, von Interesse.

Tabelle 31.3: Mittlere Verwirklichung verschiedener Formen der inneren Differenzierung nach Bildungsgang (1 = trifft gar nicht zu, 2 = trifft eher nicht zu, 3 = trifft eher zu, 4 = trifft genau zu).

	HS	RS	IGS	GY	Gesamt
Gemeinsamer Unterricht	3.5	3.6	3.4	3.6	3.6
Gezielte Zusatzaufgaben	2.7	3.0	2.5	2.6	2.8
Schnelle Schüler dürfen zum Nächsten übergehen	3.0	2.7	3.1	2.4	2.7
Verlange von guten Schülern mehr	2.9	2.7	2.9	2.5	2.7
Fähigkeitsheterogene Gruppen	2.5	2.6	2.8	2.7	2.6
Verlange von Leistungsschwachen weniger	2.6	2.5	2.7	2.2	2.4
Extraaufgaben für Leistungsstarke	2.6	2.4	2.5	2.2	2.4
Verschiedene Aufgaben bei Stillarbeit	2.4	2.2	2.6	2.1	2.2
Gruppenbildung: andere Kriterien	1.7	2.2	1.9	2.5	2.2
Fähigkeitshomogene Gruppen	1.8	1.9	2.1	1.8	1.9
Unterschiedliche Hausaufgaben	2.1	1.8	1.7	1.8	1.8

Anmerkungen: HS: Hauptschule; RS: Realschule; IGS: Integrierte Gesamtschule; GY: Gymnasium.

Die Ergebnisse in Tabelle 31.3 zeigen, dass Maßnahmen der inneren Differenzierung insbesondere durch zusätzliche Aufgaben, eine schnellere Taktung und ein höheres Erwartungsniveau gekennzeichnet sind. Gezielte Zusatzaufgaben spielen dabei in der Realschule eine signifikant größere Rolle als in den anderen Bildungsgängen, während die unterschiedliche Taktung in der Integrierten Gesamtschule überwiegt. Der Einsatz fähigkeitshomogener Gruppen und die Erteilung unterschiedlicher Hausaufgaben sind – insgesamt betrachtet – am wenigsten verbreitet.

Umgang mit Fehlern

Anders als schriftliche Leistungsüberprüfungen gehört die mündliche Korrektur von Schüleräußerungen zum täglichen Geschäft des Englischlehrers. Bei wenigen Themen der Englischdidaktik findet sich eine solche Kluft zwischen der riesigen Anzahl an Publikationen über dieses Thema und der minimalen empirischen Basis. Königs (2003) schreibt dazu: „Die Analyse des Phänomens (Fehler) führt zu dem Befund, dass wir über mündliche Fehlerkorrekturen aus empirisch abgesicherter Sicht wenig wissen“ (S. 379).

Das Thema „Fehler“ spielt bei DESI eine wichtige Rolle: Hierzu wurden sowohl Lehrpersonen als auch Schülern eine Reihe – größtenteils äquivalenter – Fragen gestellt. Darüber hinaus stehen uns detaillierte beobachtungsbasierte Informationen aus der Videostudie des Englischunterrichts zur Verfügung (vgl. Kapitel 29).

Für die DESI-Befragung haben wir für das vorliegende Kapitel aus der Vielzahl theoretisch interessanter Aspekte im Zusammenhang mit der mündlichen Korrektur von Fehlern zwei Bereiche ausgewählt: a) Vergleich verschiedener Unterrichtsphasen (sprach- vs. kommunikationsbezogen) im Hinblick auf die Häufigkeit, mit der unterschiedliche *Typen von Fehlern* korrigiert werden, und b) Formen des Umgangs

mit diesen Fehlern: Aussprache- und Grammatikfehler. Abbildung 31.2 stellt den Profilvergleich der Fehlerkorrektur in Abhängigkeit vom unterrichtlichen Kontext (sprachbezogene oder kommunikative Phase) dar. Eine dreifaktorielle Varianzanalyse (ohne Abbildung) erbringt signifikante Haupteffekte für die Unterrichtsphase, den Fehlertyp und den Bildungsgang sowie signifikante Wechselwirkungen zwischen Fehlertyp und Bildungsgang sowie Unterrichtsphase und Fehlertyp¹.

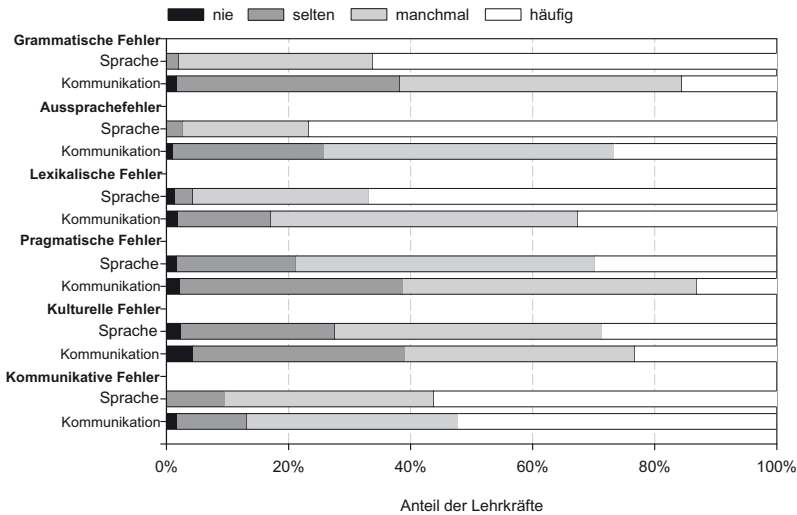


Abbildung 31.2: Korrektur verschiedener Typen von Fehlern in sprachbezogenen und kommunikativen Phasen des Englischunterrichts.

Der Ergebnistrend liegt im Sinne normativer Erwartungen vieler Fremdsprachendidaktiker, wonach die Fehlertoleranz in kommunikativen Phasen des Unterrichts – über verschiedene Fehlertypen hinweg – ausgeprägter sein sollte als in rein sprachbezogenen Phasen: „Unter dem Prinzip, dass das ‚Was‘ der Mitteilung wichtiger ist als das ‚Wie‘, wird für eine größere Toleranz im Bereich des Sprachstandards plädiert“ (Heuer/Klippel 1993). Dies entspricht der geläufigen Kurzformel „*message before accuracy*“ und umfasst insbesondere Fehler in der Aussprache und der Grammatik. Bildungsgangunterschiede zeigen sich vor allem bei lexikalischen Fehlern, die in der Hauptschule und der IGS eher toleriert werden als in den anderen Bildungsgängen.

Abbildung 31.3 präsentiert unterschiedliche Formen des Umgangs mit Fehlern aus Lehrersicht². Diese kommen unterschiedlich häufig vor (signifikanter Haupteffekt), wobei die unterschiedliche Vorkommenshäufigkeit auch von Bildungsgang und

1 Haupteffekte: Unterrichtsphase ($p < .001$), Fehlertyp ($p < .001$), Bildungsgang ($p < .01$); Wechselwirkungen: Fehlertyp x Bildungsgang ($p < .001$), Unterrichtsphase x Fehlertyp ($p < .001$).

2 Der Einbezug auch der Schülersicht hätte den Rahmen dieses Kapitels gesprengt; zudem bietet die Videostudie zur Frage des Umgangs mit Fehlern noch aussagekräftigere Ergebnisse.

Fehlerart (Aussprache-, Grammatikfehler) abhängt (signifikante Wechselwirkungen zweiter und dritter Ordnung)³.

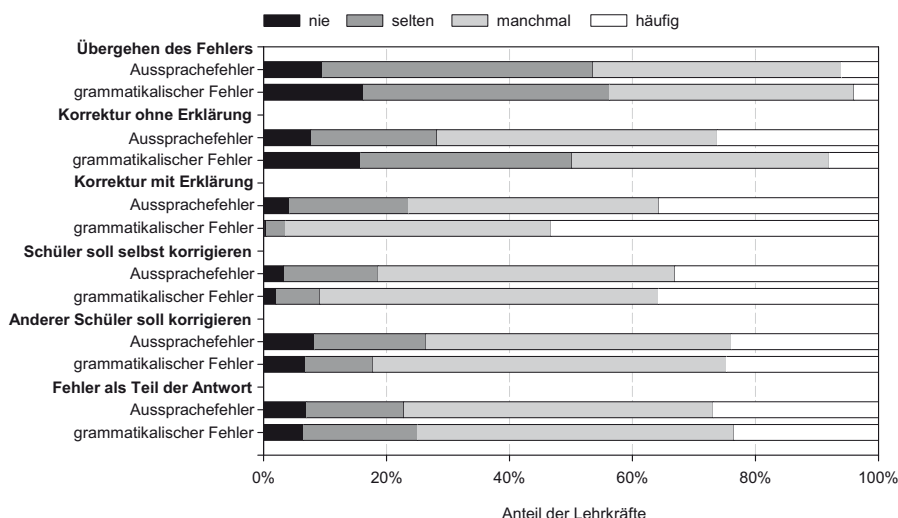


Abbildung 31.3: Formen des Umgangs mit Aussprache- und Grammatikfehlern im Englischunterricht.

Fehler zu übergehen ist aus Lehrersicht weniger üblich, als sie zu korrigieren, wobei Korrekturen *ohne Erklärung* bei Aussprachefehlern wesentlich häufiger vorgenommen werden als bei grammatischen Fehlern; bei Korrekturen *mit Erklärungen* verhält es sich genau umgekehrt. Ein didaktisch wichtiger Aspekt, die *Selbstkorrektur* von Fehlern durch Schülerinnen und Schülern, erfolgt – je nach Fehlertyp – aus Lehrersicht in 81% bis 90%, aus Schülersicht in 76% bis 77% der Fälle regelmäßig. Die Videostudie (vgl. Kapitel 29) zeichnet ein vollkommen anderes Bild der Unterrichtspraxis als die fragebogenbasierten Lehrer- und Schülerangaben: So werden Schüler tatsächlich nur in 14% aller Korrekturfälle produktiv-kognitiv involviert. Im Bereich „Selbstkorrektur“ (Lehrerangaben) finden sich auch die ausgeprägtesten Bildungsgangunterschiede: Am häufigsten erhalten dazu Schülerinnen und Schüler in Gymnasien die Gelegenheit, gefolgt von Real- und Hauptschülern.

Textsorten und Material

Um einen Überblick über die Verwendung unterschiedlicher Arten von Unterrichtsmaterial und Textsorten zu erhalten, wurden 50 Bezeichnungen von Text- und Unterrichtsmaterialien mit der Frage nach der Häufigkeit der Verwendung im Unterricht vorgegeben. Diese umfassten authentische Sachtexte (z.B. Tabellen, Wegbeschreibungen, Werbe- und Zeitungstexte), literarische Texte (z.B. Balladen,

3 Haupteffekte: Umgang mit Fehlern ($p < .001$); Wechselwirkungen: Umgang mit Fehlern x Bildungsgang ($p < .001$), Umgang mit Fehlern x Fehlerart ($p < .001$) Fehlerart x Umgang mit Fehlern x Bildungsgang ($p < .01$).

Novellen, Gedichte, Satiren) sowie Texte mit vorwiegend kommunikativ-interaktivem Bezug (z.B. Einladungen, Briefe und Dialoge). Weil die Zusammenfassung zu Skalen wichtige Informationen maskieren würde, berichten wir in Abbildung 31.4 die Ergebnisse auf der Ebene der Einzelitems.

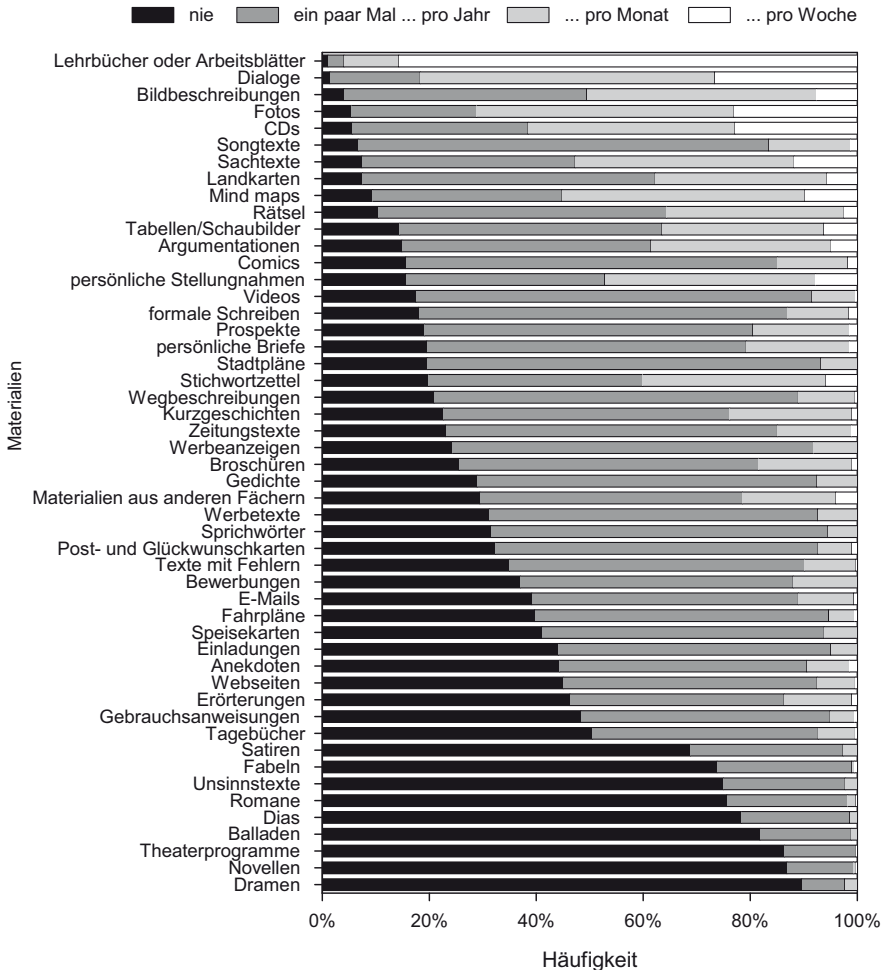


Abbildung 31.4: Textsorten und Material im Englischunterricht.

Abbildung 31.4 ist von oben nach unten in aufsteigender Reihenfolge bezüglich der Antwortkategorie „nie“ sortiert, um den Blick auf gängige und eher „exotische“ Typen von Lehr-Lern-Material zu vereinfachen. Lehrbücher und Arbeitsblätter finden sich erwartungsgemäß an der Spitze, während literarische Texte (wie Dramen, Novellen und Romane) insgesamt gesehen eine untergeordnete Rolle spielen. Eine Betrachtung der verschiedenen Materialtypen nach Bildungsgängen (ohne Abbildung) zeigt zweierlei: Zum einen ist die absolute Häufigkeit der Verwendung unterschiedlicher Arten von Material in der Hauptschule besonders gering; zum anderen ist der

relative Anteil an literarischen Texten (gemessen an allen Texten) im Gymnasium erwartungsgemäß am höchsten.

Unterrichtssprache

Zahlreich sind die Theorien und Behauptungen, ob überhaupt – und wenn ja: unter welchen Bedingungen und in welchen Kontexten – im Englischunterricht auch einmal die Erstsprache (hier: Deutsch) verwendet werden darf. Das Spektrum reicht von der aufgeklärten Einsprachigkeit (Butzkamm 1998) bis hin zu Befürwortern einer strikten Einsprachigkeit. Bei DESI wurden zu dieser Frage sowohl die Lehrpersonen als auch die Schüler/innen befragt, und zwar zur Häufigkeit des Gebrauchs der deutschen Sprache in verschiedenen Unterrichtssituationen. Warum diese Erhebung aus beiden Perspektiven? Weil es sein könnte, dass das Gebot der Einsprachigkeit mancherorts als normativ-unumstößlich und als didaktische Wahrheit angesehen wird und man deshalb nicht ausschließen kann, dass Lehrerangaben bisweilen eher den Wunsch („social desirability“) als die Wirklichkeit abbilden. Darüber hinaus verfügen wir über die beobachtungsbasierten Ergebnisse der DESI-Videostudie (vgl. Kapitel 29). Abbildung 31.5 zeigt die Ergebnisse aus Lehrersicht.

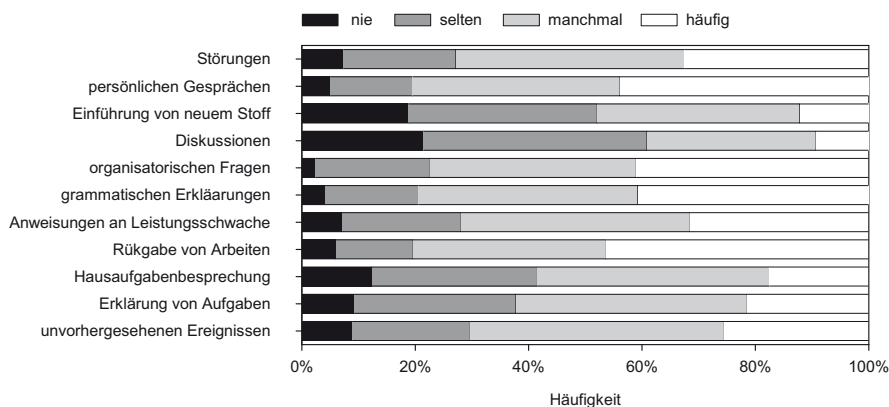


Abbildung 31.5: Verwendung der deutschen Sprache in unterschiedlichen Situationen im Englischunterricht aus Lehrersicht.

Die Ergebnisse sind weitgehend selbsterklärend: Wenn im Englischunterricht Deutsch gesprochen wird, dann am ehesten bei persönlichen Gesprächen, bei der Rückgabe von Arbeiten, organisatorischen Fragen und grammatischen Erklärungen – und am wenigsten bei Diskussionen, der Einführung von neuem Stoff und der Besprechung der Hausaufgaben. Starke Bildungsgangunterschiede (ohne Abbildung) resultieren daraus, dass Gymnasiallehrer laut Selbstauskunft insbesondere bei Diskussionen und bei der Einführung neuen Stoffs so gut wie nie die deutsche Sprache verwenden. Ein differenzierter Vergleich der Lehrer- mit der Schülerperspektive muss hier aus Platzgründen entfallen; stattdessen verweisen wir auf die einschlägigen Ergebnisse der Videostudie (Kapitel 29).

Umgang mit den Hausaufgaben

Ob und unter welchen Umständen Hausaufgaben eine nützliche oder sogar nötige Ergänzung des Schulklassenunterrichts sind, und unter welchen Bedingungen sie wirksam sind, dies ist Gegenstand zahlreicher neuerer Arbeiten in der Pädagogischen Psychologie (Trautwein/Köller/Baumert 2001). Unabhängig davon sind Hausaufgaben auch in der Fremdsprachendidaktik (Pauels 2003) ein breit diskutiertes Thema, wenngleich die empirische Evidenz hierzu bisher weitgehend fehlt. Einigkeit scheint nur in der Hinsicht zu bestehen, dass Hausaufgaben – wenn sie denn aufgegeben werden – kontrolliert werden müssen. Zum Thema Hausaufgaben wurden bei DESI Lehrerinnen und Lehrer sowie Schülerinnen und Schüler mit identischen Items befragt. Abbildung 31.6 zeigt die Ergebnisse.

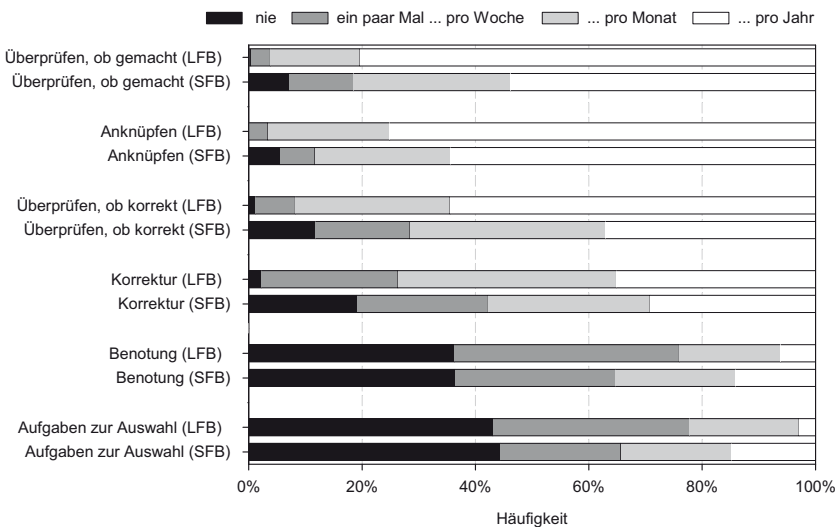


Abbildung 31.6: Umgang mit den Hausaufgaben aus Lehrersicht (Lehrerfragebogen, LFB) und Schülersicht (Schülerfragebogen, SFB).

Wie ersichtlich, scheint die Vergabe von Hausaufgaben und ihre anschließende Besprechung – nicht dagegen ihre Benotung – Teil des regulären Unterrichtsalltags zu sein. Aus Schülersicht ergibt sich teilweise ein etwas anderes Bild, insbesondere was die Häufigkeit der Überprüfung angeht (d.h. ob sie überhaupt gemacht wurden bzw. ob sie korrekt sind): der Anteil der Lehrpersonen, die angeben, dies regelmäßig (mindestens ein paar Mal pro Woche) zu tun, ist um gut 20% höher als aus Schülersicht. Drei Items (Aufgeben von Hausaufgaben sowie Besprechen und Anknüpfen) werden zur Skala „Integration der Hausaufgaben“ zusammengefasst, die einen Aspekt der Unterrichtsqualität aus Schülersicht darstellt und im Kapitel 32 („Unterrichtswirksamkeit“) thematisiert wird.

Beteiligung von Schülerinnen und Schülern

Je nach Unterrichtsfach und -inhalt und in Abhängigkeit von der Klassenstufe gibt es zahlreiche Möglichkeiten, Schülerinnen und Schüler einzubeziehen, ihnen Gestaltungsmöglichkeiten und Mitbestimmungsmöglichkeiten zu eröffnen, um zum einen die Selbständigkeit zu fördern und zum anderen Fachinteresse und Engagement zu erhöhen (Krapp 2001). In der deutschsprachigen Schulforschung gibt es seit Fend (1976) unter dem Stichwort „Autonomie“ oder „Freiräume für Selbständigkeit“ eine intensive empirische Tradition. Welche Möglichkeiten der Beteiligung sind im Englischunterricht anzutreffen? Die Ergebnisse berichtet Abbildung 31.7. Das Ergebnismuster zeigt: Schülerinnen und Schüler werden insgesamt gesehen so gut wie gar nicht – am ehesten noch im Gymnasium – an der inhaltlichen Gestaltung des Englischunterrichts beteiligt und erfahren kaum Wahlmöglichkeiten, weder bei den Hausaufgaben noch in anderen Bereichen, in denen dies prinzipiell möglich wäre.

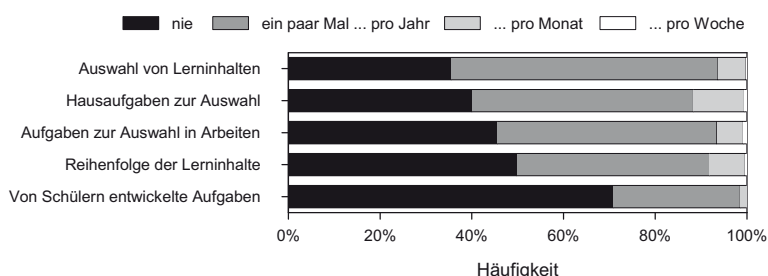


Abbildung 31.7: Beteiligung von Schülerinnen und Schülern (Lehrerangaben).

Ausblick

Das Ziel dieses Kapitels bestand darin, ausgewählte Merkmale der Alltagspraxis des Englischunterrichts zu beschreiben. Obwohl gelegentlich auf Bildungsgangunterschiede eingegangen wurde, lag der Schwerpunkt auf allgemeinen, bildungsgangübergreifenden Ergebnismustern. Aus Platzgründen mussten wir eine strikte Auswahl der bei DESI untersuchten Inhaltsbereiche des Unterrichts vornehmen. Die Analyse dieser Punkte bleibt künftigen vertiefenden Analysen vorbehalten. Dazu gehören neben bildungsgangspezifischen Auswertungen beispielsweise Aspekte der Unterrichtszeit und der Lerngelegenheiten (Unterrichtsausfall, Fehlzeiten von Schülerinnen und Schülern), der Hausaufgabenzeit sowie der systematische Vergleich von Angaben zu äquivalenten Gegenstandsbereichen aus unterschiedlichen Perspektiven (Lehrer- vs. Schülerfragebogen) und -methoden (Befragung versus Videografie).

Literatur

- Butzkamm, W. (1998): Zehn Prinzipien des Fremdsprachenlernens und -lehrens. In: Timm, J.-P.: Englisch lernen und lehren. Berlin: Cornelsen, S. 45-52.
- Fend, H./Knörzer, W./Nagl, W./Specht, W./Väth-Szud-Ziara, R. (1976): Sozialisationseffekte der Schule. Soziologie der Schule II. Weinheim: Beltz.
- Helmke, A. (2004): Unterrichtsqualität: Erfassen, Bewerten, Verbessern. Seelze: Klett-Kallmeyer.
- Helmke, A./Hosenfeld, I./Schrader, F.-W./Wagner, W. (2002): Unterricht aus der Sicht der Beteiligten. In: Helmke, A./Jäger, R. S. (Hrsg): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik, S. 325-411.
- Helmke, A./Jäger, R. S. (Hrsg.) (2002): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik.
- Heuer, H./Klippel, F. (1993): Englischmethodik. Problemfelder, Unterrichtswirklichkeit, Handlungsempfehlungen. Berlin: Cornelsen.
- Königs, F. G. (2003): Fehlerkorrektur. In: Bausch, K.-R./Christ, H./Krumm, H.-J.: Handbuch Fremdsprachenunterricht. Tübingen und Basel: A. Francke Verlag, S. 377-382.
- Krapp, A. (2001): Interesse. In: Rost, D. H.: Handwörterbuch Pädagogische Psychologie. Weinheim: Psychologie Verlags Union, S. 286-294.
- Pauels, W. (2003): Hausaufgaben. In: Bausch, K.-R./Christ, H./Krumm, H.-J.: Handbuch Fremdsprachenunterricht. Tübingen und Basel: A. Francke Verlag, S. 317-320.
- Trautwein, U./Köller, O./Baumert, J. (2001): „Lieber oft als zu viel: Hausaufgaben und die Entwicklung von Leistung und Interesse im Mathematik-Unterricht der 7. Jahrgangsstufe.“ In: Zeitschrift für Pädagogik 47, H. 5, S. 103-124.

*Andreas Helmke / Tuyet Helmke / Friedrich-Wilhelm Schrader /
Wolfgang Wagner / Eckhard Klieme / Günter Nold / Konrad Schröder*

32 Wirksamkeit des Englischunterrichts

Viele Themen der Unterrichtsqualität werden in der deutschen Englischdidaktik (ebenso wie in der internationalen Diskussion zu „English as a Foreign Language“) seit langem ausführlich und vielfach kontrovers diskutiert (Bausch/Christ/Krumm 2002; Butzkamm 1998, 2004; Klippel 2004; Timm 1998). Sondiert man aussagekräftige empirischen Studien zur Qualität des *Englischunterrichts*, dann wird man allerdings – zumindest im deutschen Sprachraum – kaum fündig: In internationalen Leistungsstudien dominieren mathematische und naturwissenschaftliche Inhalte. Zu den wenigen einschlägigen large-scale-Studien zählen die betagte IEA Study „The teaching of English as a Foreign Language“ (Lewis/Massad 1975), die Konstanzer Schulleistungsstudien der Forschungsgruppe Fend (Fend 1984, 1998; Haenisch 1979; Lukesch 1979) sowie die Studien von Achtenhagen (Wienold u.a. 1985), BIJU (Gruehn 2000), TOSCA (Köller/Trautwein 2004) und dem Schulinspektorat Graubünden (2004). Allerdings sind diese Studien mit Ausnahme von BIJU querschnittlich angelegt, so dass die Erkenntnismöglichkeiten – sofern überhaupt der Unterricht erfasst wurde – beschränkt sind.

Im Prinzip käme auch die empirische Unterrichtsforschung zum *Zweitspracherwerb* in anderen Ländern als Quelle für die Entwicklung von Instrumenten und Hypothesen in Frage. Denn für die Frage der Unterrichtsqualität ist es eigentlich nicht ausschlaggebend, ob sich der Fremdsprachenunterricht auf „English as a Foreign Language“ in Deutschland oder in Vietnam, auf die Fremdsprache Spanisch in den USA oder Deutsch in Polen bezieht. Aber auch hier ist die Forschungslage sehr eingeschränkt (Judd/Tan/Walberg 2001).

Am ergiebigsten für dieses Kapitel sind die Studien, die – mit fachdidaktischem Hintergrund – den Unterricht mit Hilfe standardisierter Verfahren, insbesondere durch *Unterrichtsbeobachtungen*, erfasst und mit abhängigen Variablen in Beziehung gebracht haben. Hierzu gibt es speziell im angloamerikanischen Raum eine große Zahl von Instrumenten und Studien (Waxman/Tharp/Hilberg 2004). Für dieses Kapitel relevant sind vor allem die Studien, die das Verfahren COLT (Communicative Orientation of Language Teaching) eingesetzt haben (Spada/Fröhlich 1995). Die dort entwickelten Kategorien bilden einen wesentlichen Teil des Fundamentes der DESI-Videostudie und regten auch die Instrumentenentwicklung für die fragebogenbasierte Hauptstudie an. Dies betrifft vor allem diejenigen Elemente, die sich auf die Förderung kommunikativer Kompetenzen von Schülerinnen und Schülern beziehen. Guter Englischunterricht wäre aus dieser Perspektive demnach wie folgt charakterisierbar:

- Längere sprachliche Schüleräußerungen anstatt von Ein-Wort-Sätzen

- Beteiligung von Schülerinnen und Schülern an der Gestaltung und den Inhalten des Unterrichts
- Verwendung von Englisch als Verkehrssprache im Unterricht
- Ermöglichung der Selbstkorrektur von Schülerfehlern
- Balance von „accuracy“ und „fluency“

Auf der Grundlage der in der DESI-Studie verwendeten Kategorien und des dort verwendeten Designs (Zwei-Punkt-Erhebung) lässt sich das in Kapitel 26 skizzierte allgemeine Angebots-Nutzungs-Modell (Rahmenmodell der Unterrichtswirksamkeit) wie in Abbildung 32.1 dargestellt spezifizieren.

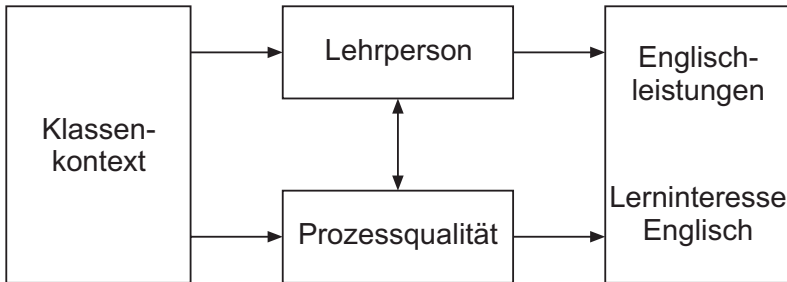


Abbildung 32.1: Spezifikation des Angebots-Nutzungs-Modells für den Englischunterricht.

Dieses Modell ist gegenüber dem umfassenderen Modell in Kapitel 26 stark vereinfacht und umfasst lediglich diejenigen Variablen, die in diesem Bericht tatsächlich analysiert werden.

Klassenkontext: Dies sind Merkmale der Klassenzusammensetzung.

Lehrperson: Ausgewählt wurden Merkmale, die einen engen Bezug zur beruflichen Tätigkeit und zur Schülerleistung erwarten lassen.

Prozessqualität: Dieser zentrale Block umfasst verschiedene, mit Hilfe dimensionsanalytischer Methoden voneinander unterscheidbarer Aspekte, die im folgenden Abschnitt detailliert beschrieben werden.

Zielkriterien: Als Zielkriterien fungieren bei den folgenden Auswertungen primär diejenigen beiden Aspekte der englischen Sprachkompetenz, zu denen von beiden Messzeitpunkten zu Beginn und am Ende der neunten Jahrgangsstufe Daten vorliegen, nämlich das Hörverstehen und die Fähigkeit zur Textrekonstruktion (C-Test), (vgl. Kapitel 11 und 14). Außerdem berücksichtigen wir das Lerninteresse im Fach Englisch (Beispielitems: „Ich finde das Fach Englisch cool“; „In meiner Freizeit beschäftige ich mich manchmal mit Englisch über die Hausaufgaben hinaus“; „Ich lege großen Wert darauf, in Englisch gut zu sein“).

32.1 Unterrichtspraxis, Lehrpersonmerkmale und Unterrichtswirksamkeit

Unterrichtspraxis aus Lehrersicht. Wir verzichten hier auf eine komplette Darstellung der Korrelationen zwischen Unterrichtsmerkmalen und Zielkriterien, weil

sich die bei den Lehrern erfragten Aspekte der Alltagspraxis für die Entwicklung der Englischkompetenzen als weitgehend irrelevant herausgestellt haben. Es zeigen sich zwar teilweise sehr hohe einfache Korrelationen des Leistungsniveaus (z.B. mit der Verwendung der deutschen Sprache: $-.67, p < .001$ mit Leistungsdifferenzierung: $-.36, p < .001$ und mit der Verwendung literarischer Texte: $.36, p < .001$). Partialisiert man aber Bildungsgang, die Ausgangsleistungen zu Beginn der neunten Jahrgangsstufe, kognitive Grundfähigkeit, Geschlechts- und Sprachzusammensetzung aus, dann verbleiben lediglich die folgenden statistisch signifikanten, aber zahlenmäßig nur geringen partiellen Korrelationen (für das Lerninteresse im Fach Englisch zeigen sich keine statistisch signifikanten Korrelationen):

- für den C-Test: Übergehen kommunikativer Fehler, $r = .03, p < .05$
- für den Hörverstehenstest: Übergehen grammatischer Fehler, $r = -.08, p < .01$

Merkmale der Expertise und Erfahrung von Lehrkräften. Ergiebiger ist die Befundlage bei den in Kapitel 30 beschriebenen Personmerkmalen. Hierzu berichten wir im Folgenden in Gestalt je eines Strukturgleichungsmodells Ergebnisse zu den beiden Zielkriterien „Hörverstehensleistung“ und „Lerninteresse im Fach Englisch“. Alle Schülermerkmale in den Abbildungen 32.2 und 33.3 beziehen sich auf die Schulklassenebene. Zugrunde liegt ein zweiebenenanalytisches Modell (wir berichten nur die Ergebnisse auf Klassenebene); auspartialisiert wurden Bildungsgang, Sprach- und Geschlechtszusammensetzung.

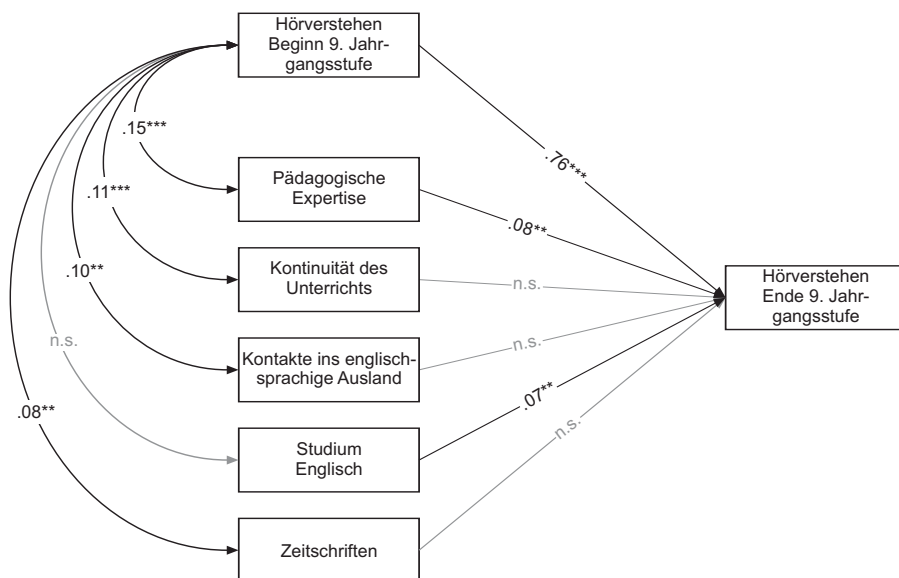


Abbildung 32.2: Lehrpersonmerkmale und Kompetenzzuwachs im Hörverstehen Englisch in der neunten Jahrgangsstufe.

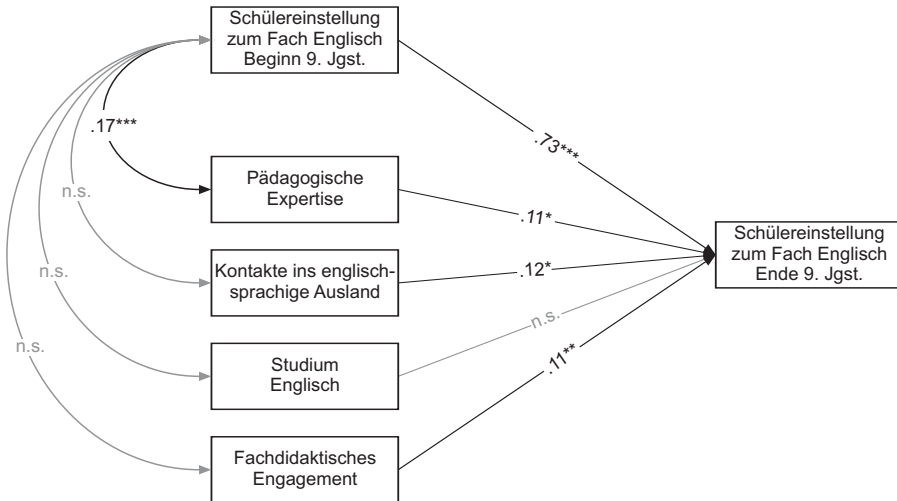


Abbildung 32.3: *Lehrpersonenmerkmale und Zuwachs des Lerninteresses im Fach Englisch in der neunten Jahrgangsstufe.*

Die Ergebnisse unterstreichen zunächst die bereits zuvor berichtete hohe Stabilität sowohl der Hörverstehensleistung als auch des Lerninteresses im Fach Englisch. Man kann es allerdings auch so sehen: Ein Pfadkoeffizient von .76 (Pretest auf Posttest beim Hörverstehen) bedeutet, dass knapp die Hälfte der Leistungsunterschiede ($.76 \times .76 = .58$) zwischen Klassen nicht festgelegt sind und somit Spielräume für unterrichtliche Einflüsse eröffnen. Lehrkräfte mit Ausbilder- oder Mentorstatus („pädagogische Expertise“) erreichen bei beiden Zielkriterien günstigere Ergebnisse. Eine positive motivationale Entwicklung ist insbesondere dann zu verzeichnen, wenn Lehrkräfte fachdidaktisches Interesse an den Tag legen (hier: Bezug einschlägiger Zeitschriften) und Kontakte ins englischsprachige Ausland pflegen. Dieses Ergebnismuster bestätigt empirisch bisher nicht untermauerte Vermutungen der allgemeinen Unterrichtsforschung (Rolle von Engagement, fachdidaktischer Expertise und von Auslandsaufenthalten; Ehrenreich 2004¹).

32.2 Beschreibung und Dimensionalität der Unterrichtsqualität aus Schülersicht

Basis der folgenden Analysen sind 15 Skalen im Schülerfragebogen (vgl. Tabelle 32.1), die unterschiedliche Facetten der Unterrichtsqualität aus Schülersicht abdecken. Sie sollen anhand der in Tabelle 32.1 dargestellten Beispielitems veranschaulicht werden.

¹ Allerdings ging es in dieser Arbeit nur um den Ertrag des Fremdsprachenassistentenjahres.

Tabelle 32.1: Skalen zur Unterrichtsqualität: Konstrukte, Beispielitems, Mittelwerte und Standardabweichungen, Reliabilitäten (Cronbachs α) und Intra-Klassen-Korrelationen (vgl. DESI-Skalenhandbuch).

Konstrukt	Beispielitem(s)	M	SD	α	IKK
Strukturiertheit	Am Ende der Stunde fasst mein Englischlehrer das Wichtigste zusammen.	2.25	0.72	.77	.15
Positive Fehlerkultur	Wenn ich im Englischunterricht Fehler mache, bespricht mein Englischlehrer diese mit mir auf eine Art und Weise, dass es mir wirklich etwas bringt.	2.80	0.55	.89	.20
Unterrichtsklima	Ich komme mit meinem Englischlehrer gut aus.	2.66	0.74	.83	.16
Motivierung thematisch	Mein Englischlehrer kann auch trockene Themen wirklich interessant machen.	2.40	0.82	.79	.20
Motivierung instrumentell	Mein Englischlehrer weist darauf hin, wie nützlich Englisch im Alltag ist.	2.79	0.75	.83	.12
Schülerorientierung	Mein Englischlehrer ermuntert mich, meine persönlichen Erfahrungen in den Unterricht einzubringen.	2.65	0.66	.87	.22
Unterstützung	Mein Englischlehrer achtet darauf, dass ich im Unterricht mitkomme.	2.80	0.70	.84	.17
Verständlichkeit	Mein Englischlehrer drückt sich klar und deutlich aus.	2.82	0.70	.76	.19
Wichtigkeit von Kommunikation	Wie wichtig ist deinem Englischlehrer, dass ihr euch auf Englisch unterhalten könnt?	3.04	0.64	.82	.19
Wichtigkeit von Korrektheit	Wie wichtig ist deinem Englischlehrer, dass eure Aussprache korrekt ist / eure Sätze beim Sprechen grammatikalisch korrekt sind?	3.16	0.62	.87	.17
Integration der Hausaufgaben	Wird im Unterricht auf den Inhalt der Hausaufgaben eingegangen?	3.61	0.66	.83	.26
Management	Mein Englischlehrer sorgt dafür, dass die Schüler die ganze Stunde über aufpassen.	2.64	0.81	.77	.22
Störungsfreiheit	Mein Englischlehrer muss lange warten, bis Ruhe eintritt.	2.75	0.69	.87	.23
Tempodruck	Wenn mein Englischlehrer mir eine Frage gestellt hat, lässt er / sie mir genug Zeit zum Überlegen.	2.03	0.76	.79	.10
Verwendung der deutschen Sprache	Denk einmal an deinen Englischunterricht in der vergangenen Zeit. Wie oft wird im Englischunterricht die deutsche Sprache verwendet? bei Störungen der Disziplin in der Klasse	2.79	0.67	.89	.29

Anmerkung: In der Tabelle wird aus Platzgründen die Formulierung „mein Englischlehrer“ verwendet; im Fragebogen ist dagegen selbstverständlich von „mein Englischlehrer / meine Englischlehrerin“ die Rede.

Zusätzlich setzten wir ausgewählte Items zum Thema „Umgang mit Schülerfehlern“ aus Schülersicht ein, die sich nicht zu Skalen zusammenfassen lassen (vgl. Kapitel 31), und zwar den Umgang mit grammatischen und mit Aussprachefehlern (*Ignorieren / Gelegenheit zur Selbstkorrektur geben / Korrigieren*).

Zur Klärung der Dimensionalität führten wir explorative Faktorenanalysen durch. Es wurden alle in Skalenform vorliegenden Schülerperzeptionen des Englischunterrichts einbezogen, mit Ausnahme der auf spezifische Kompetenzen und Testdomänen bezogenen Lerngelegenheiten. Letztere werden je nach Zielkriterium als weitere Erklärungsvariable in die Modelle einbezogen. Tabelle 32.2 berichtet die Interkorrelationen der sechs Merkmale.

Tabelle 32.2: Signifikante Interkorrelationen der auf Schülerangaben basierenden Unterrichtsmerkmale². Nicht signifikante Korrelationen ($p \geq .05$) sind nicht dargestellt.

Unterrichtsmerkmale	Merkmale				
	(2)	(3)	(4)	(5)	(6)
(1) Prozessqualität Positive Fehlerkultur Unterrichtsklima Motivierung thematisch Schülerorientierung Unterstützung Verständlichkeit	.76***	.73***	-.54***	-.27**	.68***
(2) Anspruchsniveau Wichtigkeit von Kommunikation Wichtigkeit von Korrektheit Motivierung instrumentell Integration der Hausaufgaben		.72***	-.61***	-.47***	.26**
(3) Klassenführung Management Störungsfreiheit			-.40**	-.40**	.46***
(4) Tempodruck				.27**	
(5) Verwendung der deutschen Sprache					
(6) Strukturiertheit					

Anmerkungen: Signifikanzniveaus: * $p < .05$, ** $p < .01$, *** $p < .001$ (zweiseitige Tests).

Es ergeben sich sechs Faktoren (die zugehörigen Subskalen tauchen unter dem fett gedruckten Namen des Faktors auf), wobei nur die drei ersten Faktoren mehrere Skalen umfassen: Prozessqualität, Anspruchsniveau und Effizienz der Klassenführung. Die drei restlichen Faktoren sind identisch mit einzelnen Skalen (Tempodruck, Verwendung der deutschen Sprache und Strukturiertheit).

32.3 Zur Rolle von Schulklassenunterschieden

Analysen zur Wirksamkeit des Unterrichts sind nur dann aussichtsreich, wenn sich Schulklassen im Hinblick auf die erfassten Unterrichtsmerkmale wie auch die

2 Die konfirmatorische Faktorenanalyse erfolgte in einem Zwei-Ebenen-Modell, wobei Faktoren lediglich auf der Klassenebene modelliert wurden. Auf der Individualebene wurden die Korrelationen sämtlicher Skalen untereinander modelliert. Modellfit: $CFI = .985$; $TLI = .958$; $RMSEA = .023$; $SRMR_{between} = .096$.

Zielkriterien überhaupt hinlänglich voneinander unterscheiden. Das Ausmaß der Schulklassenunterschiede lässt sich durch die *Intraklassen-Korrelation* charakterisieren. Dieses Maß setzt die Varianz zwischen Klassen zu der Gesamtvarianz in Beziehung. Ein hoher IKK-Wert bedeutet, dass sich ein großer Prozentsatz der Gesamtvarianz auf Unterschiede zwischen Klassen zurückführen lässt, mit anderen Worten: dass es bei dem entsprechenden Merkmal große Unterschiede zwischen Klassen gibt, deren Ursache in Unterschieden in der Klassenzusammensetzung, der Unterrichtsqualität oder in beiden liegen können. Wegen der Konfundierung von Leistungsniveau und Bildungsgang erfolgt die Analyse global (bildungsgang-übergreifend) sowie separat für die vier Bildungsgänge.

Tabelle 32.3: Englischleistung und Leistungszuwächse: Unterschiede zwischen Klassen („Intraklassen-Korrelationen“) nach Bildungsgang.

Leistung: Niveau	9. Jgst.	HS	RS	IGS	GY	Gesamt
C-Test	Beginn	.23	.37	.55	.28	.65
	Ende	.24	.33	.50	.29	.64
Hörverstehen	Beginn	.36	.55	.64	.46	.71
	Ende	.40	.57	.62	.48	.71
Leseverstehen	Ende	.19	.37	.48	.32	.58
	Beginn	.22	.39	.65	.28	.67
Soziopragmatik	Ende	.23	.44	.52	.40	.72
Grammatik	Ende	.28	.39	.60	.34	.68
Textproduktion	Ende	.29	.48	.67	.40	.77
Englischleistung (global)	Ende					
Leistung: Zuwachs						
C-Test		.12	.08	.30	.09	.12
Hörverstehen		.59	.50	.36	.65	.58
Lerninteresse						
Lerninteresse	Beginn	.04	.05	.05	.07	.08
	Ende	.03	.06	.05	.09	.09
Zuwachs		.03	.03	.08	.03	.03

Die in Tabelle 32.3 berichteten Ergebnisse zeigen: Während sich Schulklassen – abgesehen vom Bildungsgang IGS – hinsichtlich des Zuwachses beim *C-Test* nur wenig voneinander unterscheiden, gibt es erhebliche Klassenunterschiede beim Zuwachs für das *Hörverstehen*. Das Erklärungspotenzial des Unterrichts und von Merkmalen der Klassenzusammensetzung hängt also deutlich von dem jeweils betrachteten Zielkriterium ab. Bei einem integrativen und globalen Instrument wie dem *C-Test*, dessen erfolgreiche Bearbeitung ein breites Spektrum unterschiedlicher Kompetenzen erfordert (insbesondere Wortschatz, Grammatik und Lesestrategien), spielen Unterschiede in der Unterrichtsqualität offensichtlich eine geringere Rolle als beim Hörverstehen. Beim Lerninteresse sind die Klassenunterschiede deutlich geringer (ohne Tabelle). Die folgenden Analysen werden sich deshalb schwerpunktmäßig

mit der Rolle der Unterrichtsqualität für den Kompetenzzuwachs beim Hörverstehen beschäftigen.

32.4 Unterrichtsqualität und Englischleistung

Welche Rolle spielen Merkmale der Unterrichtsqualität aus Schülersicht für die Entwicklung sprachlicher Kompetenzen in der neunten Jahrgangsstufe? Die Merkmale des Unterrichts in Tabelle 32.4 entsprechen der Anordnung in Tabelle 32.2.

Tabelle 32.4: Korrelationen der Unterrichtsvariablen mit dem C-Test und dem Hörverstehenstest am Ende der neunten Jahrgangsstufe (Klassenebene). Nicht signifikante Korrelationen ($p \geq .05$) sind nicht dargestellt.

	C-Test			HV-Test		
	einfache Korrelation	Partialkorrelation mit dem Posttest	Partialkorrelation mit dem Zuwachs	einfache Korrelation	Partialkorrelation mit dem Posttest	Partialkorrelation mit dem Zuwachs
1 Prozessqualität						
Positive Fehlerkultur	.24***			.25***	.07**	.17**
Unterrichtsklima					.06*	.13*
Motivierung thematisch					.05*	.12*
Schülerorientierung	.23***	.02*	.12*	.19***		
Unterstützung						
Verständlichkeit	.26***	.02*	.14*	.22***	.06**	.15**
2 Anspruchsniveau						
Wichtigkeit: Kommunikation	.63***			.64***	.08**	.19**
Wichtigkeit: Korrektheit	.58***			.58***	.08**	.20**
Motivierung instrumentell	.23***	.04**	.21**	.25***	.06*	.15*
Integration der Hausaufgaben	.48***			.50***	.07**	.16**
3 Klassenführung						
Management	.15*			.18**	.06*	.14*
Störungsfreiheit	.35***			.34***	.08**	.18**
4 Tempodruck	-.60***			-.52***		
5 Verwendung deutscher Sprache	-.44***			-.53***	-.09***	-.22***
6 Strukturiertheit	-.39***			-.40***		

Anmerkung: Mehrebenenkorrelationen mit Mplus; Kontrollvariablen: Leistung zu Beginn der neunten Jahrgangsstufe, Geschlecht, Erstsprache, kognitive Grundfähigkeit, Bildungsgang. Signifikanzniveaus (jeweils einseitiger Test): * $p < .05$, ** $p < .01$, *** $p < .001$.

Die Ergebnisse machen deutlich:

a) Erwartungsgemäß gibt es markante Unterschiede zwischen den unbereinigten und den bereinigten Korrelationen. Beispiel „Tempodruck“: In leistungstärkeren Klassen (dies gilt sowohl für den C-Test als auch für das Hörverstehen) kommt Überforderung durch unangemessen hohe Geschwindigkeit seltener vor (Korrelationen in Höhe von $-.60$ und $-.52$). Dagegen leistet der Tempodruck keinen Erklärungsbeitrag für Klassenunterschiede im Kompetenzzuwachs im Laufe der neunten Klassenstufe.

b) Beim *C-Test* finden sich lediglich drei Unterrichtsmerkmale, denen nach Bereinigung um Bildungsgang- und Kontextunterschiede überhaupt ein signifikanter Erklärungswert zukommt: Motivierung, Verständlichkeit und Schülerorientierung. Dies war in Anbetracht der vergleichsweise geringen Klassenunterschiede im Zuwachs (vgl. Tabelle 32.3) zu erwarten gewesen. Bei der Suche nach unterrichtlichen Bedingungen des Zuwachses beim *Hörverstehen* wird man dagegen in wesentlich höherem Maße fündig. Mit Ausnahme des Tempodrucks, der Strukturiertheit sowie der Schülerorientierung und der Unterstützung sind bei allen Unterrichtsmerkmalen signifikante Ergebnisse zu verzeichnen.

32.5 Bildungsgangprofile der Unterrichtsqualität

Welche Rolle spielen Bildungsgangunterschiede und Merkmale der Schülerschaft für die Qualität des Unterrichts? Dieser Frage soll in zwei Schritten nachgegangen werden: Zunächst werden die vorgefundenen Bildungsgangunterschiede hinsichtlich der sechs zentralen Merkmale der Unterrichtsqualität beschrieben.

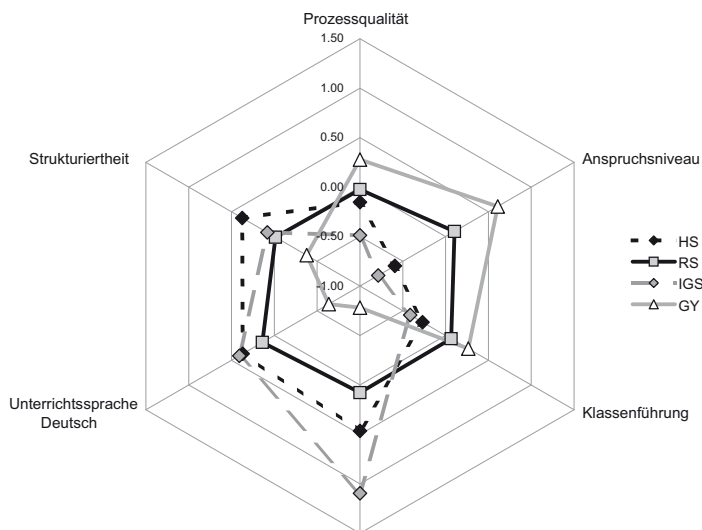


Abbildung 32.4: Bildungsgangunterschiede in den sechs zentralen Merkmalen der Qualität des Englischunterrichts aus Schülersicht.

Hier zeigt Abbildung 32.4, dass Schülerinnen und Schüler, die den Hauptschulabschluss erwerben wollen, und mehr noch Schüler/innen aus Integrierten Gesamtschulen übereinstimmend einen hohen Tempodruck sowie einen stark strukturierten Unterricht berichten. Erwartungsgemäß wird das Anspruchsniveau in Gymnasien am höchsten eingeschätzt, ebenso die Prozessqualität des Unterrichts sowie die Klassenführung.

In einem zweiten Schritt wird dem Sachverhalt Rechnung getragen, dass sich der Unterricht je nach Bildungsgang mit unterschiedlichen Eingangsvoraussetzungen konfrontiert sieht, und zwar in Bezug auf das kognitive Fähigkeits- und Vorkenntnisniveau, den Schüleranteil mit Deutsch als Erstsprache und den Mädchenanteil innerhalb der Klasse. Kontrolliert man die Bildungsgangunterschiede um diese Merkmale, dann erhält man den „Netto-Effekt“ der Bildungsgangunterschiede, d.h. Aussagen darüber, welchen Erklärungswert der Bildungsgang über die jeweils unterschiedlichen Hintergrundmerkmale hinaus für Unterschiede in der Unterrichtsqualität besitzt.

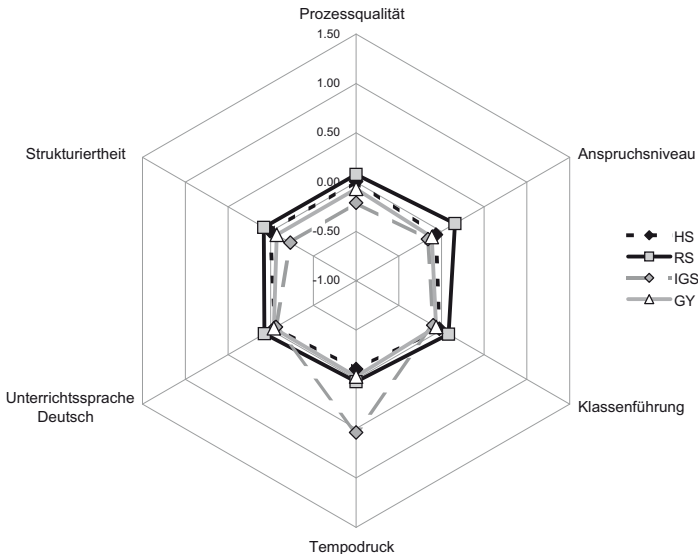


Abbildung 32.5: Um Unterschiede im Kontext (Klassenzusammensetzung) bereinigte Bildungsgangunterschiede in den sechs zentralen Merkmalen der Qualität des Englischunterrichts aus Schülersicht.

Wie Abbildung 32.5 dokumentiert, fallen mit einer Ausnahme (dem Tempodruck in der IGS) die Muster der vorher stark abweichenden Bildungsgangprofile nahezu zusammen. Dies lässt sich so interpretieren, dass der auf den ersten Blick starke Bildungsgangunterschied zwischen den Unterrichtsprofilen weitgehend durch Merkmale der je nach Bildungsgang unterschiedlich zusammengesetzten Schülerschaft vermittelt wird.

32.6 Unterrichtsqualität und Englischleistung: Komplexe Modellierung

Wir verlassen jetzt die Ebene einfacher Zusammenhänge und analysieren das Zusammenwirken von Kontext, Unterrichtsqualität und den Englischleistungen im Nachtest. Da es sehr unterschiedliche Strategien sowohl der kausalen Modellierung als auch der Berücksichtigung des Mehrebenencharakters (Individual- und Klassenebene) gibt und die folgenden Analysen den Kern der Aussagen zur Unterrichtswirksamkeit darstellen (vgl. Kapitel 26), erscheint es uns nötig, die Logik und Technik des hier verwendeten Verfahrens kurz zu skizzieren. Sofern nicht ausdrücklich etwas anderes gesagt wird, verwenden wir das Statistikprogramm *Mplus* (vgl. Muthén/Muthén 2003).

Abbildung 32.6 stellt die Hörverstehensleistung am Ende der neunten Jahrgangsstufe in Abhängigkeit von der Unterrichtssprache und den Eingangsleistungen zu Beginn der neunten Jahrgangsstufe auf Schulklassenebene dar. Aus diesen Zusammenhängen wurden die Effekte der Unterschiede von Bildungsgang, Erstsprache, kognitiver Grundfähigkeit und Mädchenanteil herausgerechnet.

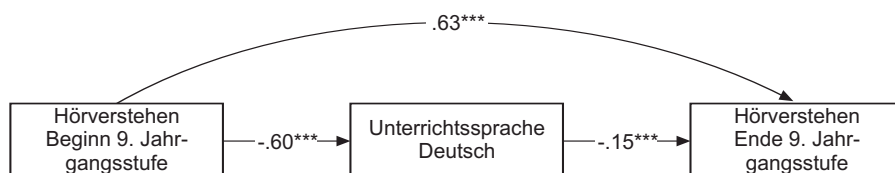


Abbildung 32.6: Direkte und indirekte – über das Merkmal „Unterrichtssprache Deutsch“ vermittelte – Wirksamkeit der anfänglichen Hörverstehenskompetenz (unter Kontrolle der Kontextmerkmale Bildungsgang, kognitive Grundfähigkeit, Mädchenanteil und Anteil von Schülern mit Deutsch als Erstsprache).

Inhaltlich bedeutet dieses Ergebnis, dass die Hörverstehensleistung von Schulklassen *erstens* äußerst stabil ist („Positionsstabilität“): Klassen, die zu Beginn der neunten Jahrgangsstufe zu den Besten gehören, sind mit hoher Wahrscheinlichkeit auch am Ende der neunten Jahrgangsstufe unter den Besten zu finden ($\beta = .63$). *Zweitens* ist die Hörverstehensleistung am Ende der neunten Jahrgangsstufe um so geringer, je häufiger im Unterricht Deutsch gesprochen wird ($\beta = -.15$). *Drittens*: Ob im Englischunterricht Deutsch oder Englisch gesprochen wird, hängt aber in erheblichem Maße vom Eingangsniveau ab: Je höher die Eingangsleistung, umso weniger wird Deutsch gesprochen ($\beta = -.60$), wobei Bildungsgangunterschiede wohlgermerkt herausgerechnet sind!

32.7 Unterrichtsqualität und Lerninteresse

Der Erfolg eines Unterrichts bemisst sich nicht nur nach dem Leistungs- oder Kompetenzzuwachs. Wir haben exemplarisch für die nicht-kognitiven Zielkriterien das Lerninteresse für das Fach Englisch ausgewählt, das zu Beginn und am Ende

der neunten Jahrgangsstufe erhoben wurde. Tabelle 32.5 berichtet zunächst die Zusammenhänge mit den Unterrichtsvariablen.

Tabelle 32.5: Korrelationen der Unterrichtsvariablen mit dem Lerninteresse Englisch am Ende der neunten Jahrgangsstufe. Nicht signifikante Korrelationen ($p \geq .05$) sind nicht dargestellt.

	einfache Korrelation	Partialkorrela- tion mit dem Lerninteresse Ende 9. Jg.St.	Partialkorrela- tion mit dem Zuwachs
1 Prozessqualität			
Positive Fehlerkultur	.68***	.22**	.43**
Unterrichtsklima	.56***	.25***	.47***
Thematische Motivierung	.62***	.30***	.58***
Schülerorientierung	.62***	.24***	.45***
Unterstützung	.56***	.23**	.45**
Verständlichkeit	.70***	.21**	.41**
2 Anspruchsniveau			
Wichtigkeit: Kommunikation	.73***		
Wichtigkeit: Korrektheit	.53***		
Instrumentelle Motivierung	.47***	.17*	.32*
Integration der Hausaufgaben	.40***		
3 Klassenführung			
Management	.44***	.13*	.25*
Störungsfreiheit	.44***	.14*	.27*
4 Tempodruck	-.59***		
5 Verwendung der deutschen Sprache	-.49***		
6 Strukturiertheit	.18*	.15*	.29*

*Anmerkung: Mehrebenenkorrelationen mit Mplus; Kontrollvariablen: Einstellung zum Fach Englisch zum Beginn der neunten Jahrgangsstufe, Geschlecht, Erstsprache, kognitive Grundfähigkeit, Bildungsgang; Signifikanzniveaus (jeweils einseitiger Test): * $p < .05$, ** $p < .01$, *** $p < .001$.*

Trotz der in Tabelle 32.3 geschilderten relativ geringen Klassenunterschiede (Intraklassen-Korrelationen) im Lerninteresse und dessen Zuwachs sind die Zusammenhänge zwischen Unterrichtsqualität und Entwicklung des Lerninteresses insgesamt noch höher als bei den Testleistungen. Einer der Gründe dafür könnte sein, dass hier Prädiktor und Zielkriterium der gleichen Quelle – dem Schülerfragebogen – entstammen. Vor allem die Komponenten des ersten Faktors, der klassische Merkmale der Unterrichtsqualität umfasst, leisten einen beachtlichen Beitrag zur Vorhersage von Unterschieden in der Entwicklung des Lerninteresses.

Vergleicht man die Ergebnismuster in den Tabellen 32.4 und 32.5 miteinander, dann wird deutlich: Die Frage, was „guten Unterricht“ ausmacht, ist je nach betrachtetem Zielkriterium unterschiedlich zu beantworten. Es gibt einen Satz von Merkmalen der Prozessqualität, die gleichermaßen für die Entwicklung des

Hörverstehens und des Lerninteresses relevant sind, es zeigen sich jedoch auch kriterienspezifische Unterschiede, beispielsweise bei der Strukturiertheit.

32.8 Klassengröße und Heterogenität der Vorkenntnisse

Wenn bisher von relevanten Kontextmerkmalen die Rede war, dann dienten diese im Wesentlichen der „Bereinigung“ der inhaltlich interessierenden Pfade (insbesondere vom Unterricht auf die Leistung) durch zentrale Kontrollvariablen. Im Folgenden möchten wir zwei weitere Merkmale untersuchen, deren Bedeutung für den Unterricht und die Leistung in der Bildungspolitik diskutiert wird: die Klassengröße und die Heterogenität der (klasseninternen) Leistungen.

Die meisten bisherigen Studien haben keinen Zusammenhang zwischen *Klassengröße* und Testleistung nachweisen können. Allerdings handelte es sich a) zu meist um Querschnittsstudien, wurde b) der Unterricht nicht oder nur am Rande thematisiert, und ging es c) überwiegend um mathematisch-naturwissenschaftlichen Unterricht. Im Fremdsprachenunterricht, bei dem die Kommunikation eine überragende Rolle spielt, könnten sich die Dinge anders darstellen. Unsere Hypothese ist deshalb, dass zahlenmäßig kleinere Klassen – alle anderen Kontextfaktoren konstant gehalten – eine günstige Bedingung für guten Unterricht und damit mittelbar auch für den Leistungszuwachs sind.

Zweitens überprüfen wir, ob es zutrifft, dass die *Heterogenität der Vorkenntnisse* zu Beginn der neunten Jahrgangsstufe eine Rolle für den Leistungszuwachs spielt, ob also z.B. eine hohe Leistungsstreuung – alle anderen Kontextfaktoren wiederum konstant gehalten – einen Risikofaktor für die Leistungsentwicklung darstellt.

Entgegen vielfach gefundenen Ergebnissen zum Mathematikunterricht – beispielsweise im Projekt MARKUS (Helmke/Jäger 2002) – zeigt sich hier durchaus ein klares Muster: Die Klassengröße geht eindeutig mit einem ungünstigeren Unterrichtsprofil einher (vgl. Abbildung 32.7). Das um die Hörverstehensleistung (exemplarisch für zwei Unterrichtsmerkmale) erweiterte Pfadmodell (ohne Abbildung) unterstützt die Annahme einer indirekten Wirkungsweise der Klassengröße: Je kleiner die Klasse, desto verständlicher wird der Unterricht wahrgenommen und desto häufiger ist Englisch die Unterrichtssprache. Diese beiden Unterrichtsvariablen wiederum wirken sich auf die Hörverstehensleistung aus.

Wirkt sich die Heterogenität der Vorkenntnisse auf den Unterricht und den Leistungszuwachs aus? Um dieser Frage nachzugehen, haben wir in einem Zweiebenenmodell den Zusammenhang zwischen der Heterogenität der Englischleistungen zu Beginn der neunten Jahrgangsstufe und dem Leistungszuwachs berechnet, unter Einbezug der folgenden Kontext- und Hintergrundmerkmale: Bildungsgang (GY vs. andere, RS vs. andere), Sozialschichtniveau, Mädchenanteil sowie Sprachherkunft. Die Ergebnisse belegen, dass sich die Heterogenität *nicht* negativ auf den Leistungszuwachs auswirkt: die Pfadkoeffizienten betragen für den Zuwachs beim Hörverstehen $-.01$, beim Zuwachs der C-Testleistung $.01$ und sind beide nicht signifikant (ohne Abbildung).

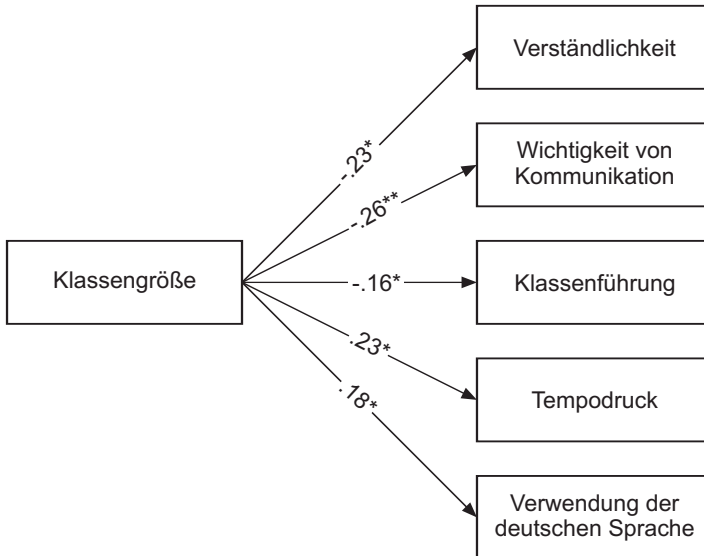


Abbildung 32.7: Klassengröße und Prozessqualität des Unterrichts (unter Kontrolle von Sozialstatus, kognitiver Grundfähigkeit, Mädchenanteil, Erstsprache, Bildungsgänge).

32.9 Ausblick

Wir haben in diesem Kapitel versucht, ein Bild der Wirksamkeit des Englischunterrichts zu zeichnen, und haben dabei drei Schülermerkmale in den Blick genommen, von denen Messungen zu Beginn und am Ende der neunten Jahrgangsstufe vorliegen, so dass Entwicklungen festgestellt werden können: Kompetenz im Bereich des *Hörverstehens* und der *Textrekonstruktion* (C-Test) sowie – als nicht-kognitives Merkmal – *Lerninteresse* im Fach Englisch. Die Ergebnisse haben gezeigt, dass die verschiedenen Zielkriterien mit unterschiedlichen Unterrichtsmustern verknüpft sind, dass es sich also gelohnt hat, anstelle eines Globalwertes „der“ Englischleistung nach unterschiedlichen Kompetenzen zu differenzieren. Zwar ist der Erklärungsbeitrag von Person- und Unterrichtsmerkmalen zahlenmäßig gering – angesichts der extrem hohen Stabilität der betrachteten Schülermerkmale jedoch beachtenswert. Insgesamt gesehen kommt Unterschieden zwischen Lehrpersonen und Unterschieden in der Qualität des Englischunterrichts eine erheblich höhere Bedeutung zu, als dies in den großen deutschen Leistungsstudien der 80er Jahre von Fend (1984) belegt werden konnte.

Wir betrachten die hier berichteten Ergebnisse als Basis, die durch weitere Analysen zunehmend ergänzt und verfeinert werden sollen.

a) Dies betrifft zunächst die Präzisierung durch die Ergebnisse der Englisch-Videostudie (vgl. Kapitel 29) und die Konkretisierung für den Vergleich von Schulen

mit bilinguaem Sachfachunterricht mit Schulen mit regulärem Sachfachunterricht (vgl. Kapitel 35).

b) Während der Schwerpunkt dieses Kapitels auf bildungsgangübergreifenden Ergebnissen lag, werden künftige Analysen bestimmten Fragen verstärkt bildungsgangspezifisch, z.B. nur für die Hauptschule, nachgehen.

d) Komplexe unterrichtliche Bereiche wie der Umgang mit *Schülerfehlern* aus Schüler- und Lehrersicht (sowie auch videobasiert), deren Berücksichtigung den Rahmen dieses Kapitels gesprengt hätten, sind späteren Publikationen vorbehalten.

c) Die Auswertungsstrategie sieht weiterhin komplexe Analysen vor, die hier aus Platzgründen nicht entfaltet werden konnten, beispielsweise um die Frage nach *Mediationsvariablen*, die die Wirksamkeit einzelner Unterrichtsmerkmale erklären, um die Vereinbarkeit unterschiedlicher *Zielkriterien* des Unterrichts, beispielsweise von Kompetenzzuwachs (von Klassen) und Verringerung der Streuung innerhalb der Klassen, um vertiefende Mehrebenenanalysen der *Schülerwahrnehmungen* des Unterrichts sowie um die Verfeinerung des Erklärungsmodells durch Einbezug weiterer Kontext- und Unterrichtsmerkmale, beispielsweise des hier nicht berücksichtigten *Klassenklimas*.

Literatur

- Bausch, K. R./Christ, H./Krumm, H. J. (Hrsg.) (2002): Handbuch Fremdsprachenunterricht. Tübingen: Francke.
- Butzkamm, W. (1998): Zehn Prinzipien des Fremdsprachenlernens und -lehrens. In: Timm, J.-P. (Hrsg.): Englisch lernen und lehren. Berlin: Cornelsen, S. 45-52.
- Butzkamm, W. (2004): Lust zum Lehren, Lust zum Lernen. Eine neue Methodik für den Fremdsprachenunterricht. Tübingen: Francke Verlag.
- Ehrenreich, S. (2004): Auslandsaufenthalt und Fremdsprachenlehrerbildung (Münchener Arbeiten zur Fremdsprachen-Forschung). München: Langenscheidt.
- Fend, H. (1984): Determinanten von Schulleistungen. Wie wichtig sind die Lehrer? In: Unterrichtswissenschaft 12, H.1, S. 62-87.
- Fend, H. (1998). Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim: Juventa.
- Gruehn, S. (2000): Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Münster: Waxmann.
- Haenisch, H. (1979): Schulleistungsvergleiche zwischen Gesamtschulen und Schularten des gegliederten Schulsystems in Nordrhein-Westfalen am Ende des 6. Schuljahres. In: Haenisch, H./Lukesch, H./Klaghofer, R./Krüger-Haenisch, E.-M. (Hrsg.): Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik. Paderborn: Schöningh, S. 6-226.
- Helmke, A./Jäger, R. S. (Hrsg.) (2002): Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik.
- Judd, E. L./Tan, L./Walberg, H. J. (2001): Teaching additional languages (Educational Practices Series, Vol. 6). Brussels: International Academy of Education & International Bureau of Education. URL: <http://www.ibe.unesco.org>.
- Klippel, F. (2004): Fachdidaktik Englisch. In: Blömeke, S./Reinhold, P. / Tulodziecki, G. /Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Braunschweig: Klinkhardt, S. 439-442.

- Köller, O./Trautwein, U. (2004): Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In: Köller, O./ Watermann, R./ Trautwein, U./Lüdtke, O. (Hrsg.): Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. Opladen: Leske + Budrich, S. 285-326.
- Lewis, E. G./Massad, C. E. (1975): The teaching of English as a foreign language in ten countries. New York, London, Sydney, Tokyo: John Wiley & Sons.
- Lukesch, H. (1979): Leistungsvergleich zwischen Gesamtschulen und herkömmlichen Schulen am Ende der Pflichtschulzeit in Nordrhein-Westfalen. In: Haenisch H./ Lukesch, H./Klaghofer, R./ Krüger-Haenisch, E.-M. (Hrsg.): Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik. Paderborn: Schöningh, S. 227-364.
- Muthén, L. K./Muthén, B. O. (2003): Mplus User's Guide. Los Angeles: Muthen & Muthen.
- Schulinspektorat des Kantons Graubünden (2004): Erhebung Pflichtfremdsprache Englisch 1. Klasse Oberstufe Schuljahr 2003/04. Schul- und Kindergarteninspektorat des Kantons Graubünden, Schweiz.
- Spada, N./Fröhlich, M. (1995): The communicative orientation of language teaching (COLT) observation scheme: coding conventions and applications. Sydney: National Centre for English Language Teaching and Research.
- Timm, J.-P. (Hrsg.) (1998): Englisch lernen und lehren. Berlin: Cornelsen.
- Waxman, H. C./Tharp, R. G./Hilberg, R. S. (Hrsg.) (2004): Observational Research in U.S. Classrooms. New Approaches for Understanding Cultural and Linguistic Diversity. Cambridge: Cambridge University Press.
- Wienold, G./Achtenhagen, F./Van Buer, J./Oldenburger, H.-A./Rösner, H./Schluroff, M. (1985): Lehrerverhalten und Lernmaterial in institutionalisierten Lehr-Lern-Prozessen. Am Beispiel des Englischanfangsunterrichts (Band 9 – 2). Göttingen: Seminar für Wirtschaftspädagogik der Georg-August-Universität.

Kerstin Göbel / Hermann-Günter Hesse

33 Vermittlung interkultureller Kompetenzen im Englischunterricht

33.1 Einleitung

Die zentrale Stellung des Faches Englisch bei der Vermittlung von Fremdsprachen bedingt, dass ihm die Aufgabe zufällt, Lernmöglichkeiten zur Entwicklung interkultureller Kompetenz bereit zu stellen. Interkulturelles Lernen gilt in Deutschland inzwischen als Kernelement des Englischunterrichts (Sekretariat KMK 1994). Dieses Unterrichtsziel wird dementsprechend in den Lehrplänen für das Fach Englisch betont (Göbel/Hesse 2004): fokussiert werden hier vor allem kulturallgemeine Lehrziele, wie beispielsweise die Überwindung von Ethnozentrismus, die Entwicklung kultureller Selbstbewusstheit, der Wertschätzung, des Interesses und des Respekts für kulturelle Verschiedenheit. Mit welchen konkreten Lehrinhalten die Lehrpersonen die allgemeinen Lehrziele vermitteln, wird in den meisten Lehrplänen nicht geregelt. Ein didaktisch sinnvolles Verhältnis von kulturspezifischen und kulturallgemeinen Aspekten interkulturellen Lernens wird in den Lehrplänen nicht erläutert. Auch die Berücksichtigung einer etwaigen *Abfolge* von interkulturellen Kompetenzstufen bei den Schülerinnen und Schülern, wie sie bspw. Bennett, Bennett und Allen (1999) vorschlagen, hat in die Konzeption von Lehrplänen bislang keinen Eingang gefunden.

Die Bildungsstandards Fremdsprache für die neunte Jahrgangsstufe (KMK 2003) gehen ausführlich auf das Lernziel interkulturelle Kompetenz ein (Steinert/Klieme 2004). Es wird definiert als Kompetenz im Umgang mit Differenz und Andersartigkeit, Verfügbarkeit von soziokulturellem Orientierungswissen und Kompetenz in der Bewältigung interkultureller Situationen (Steinert/Klieme 2004). Der für die internationale Verankerung des Fremdsprachenlernens zunehmend bedeutsame Gemeinsame Europäische Referenzrahmen (GERS) für Sprachen (Trim u.a. 2001) beschränkt seine Definition interkultureller Kompetenz bislang nur auf den sprachlichen Umgang mit Redewendungen, Aussprüchen, Registerunterschieden und sozialen Varietäten (Vollmer 2001). Fremdsprachendidaktische Konzepte über Ziele und Art der Vermittlung interkultureller Inhalte sind auch in der angloamerikanischen Literatur sehr heterogen (Paige u.a. 1999; Smith/Paige/Steglitz 1998; Bennett/Bennett/Allen 1999). Über die Umsetzung dieses Unterrichtsziels im Englischunterricht gibt es bislang nur wenig empirische Forschung (siehe Bennett/Bennett/Allen 1999; Göbel/Hesse 2004; Baron 2002; Burwitz-Meltzer 2003; Byram 1991).

Aus dem zuvor Gesagten leitet sich ab, dass bislang nicht von einem klar definierten Curriculum interkulturellen Englischunterrichts ausgegangen werden kann. Die aktuelle Situation von Englischlehrkräften in Deutschland ist weiterhin von der Tatsache gekennzeichnet, dass es kaum Lehrbücher zur Vermittlung interkultureller Kompetenz gibt, auf die sie sich in ihrem Unterricht stützen können (Baron 2002).

Ob sich Englischlehrpersonen überhaupt als Kulturvermittler sehen, und inwieweit sie interkulturellen Themen in ihrem Unterricht überhaupt eine Bedeutung einräumen, ist empirisch nicht geklärt (Byram 1991). Interkultureller Unterricht ist demnach sehr stark auf Ideen und Initiativen einzelner Lehrkräfte angewiesen.

Für die Realisierung interkulturellen Fremdsprachenunterrichts ist es erforderlich, dass Lehrpersonen über Wissen über die andere Kultur, über kulturelle Phänomene im Allgemeinen sowie Wissen über den kulturellen Hintergrund der Schülerinnen und Schüler verfügen (Edelhoff 1987; Damen 1987). Weiterhin hat die Einstellung der Lehrpersonen eine sehr große Bedeutung: Lehrpersonen sollten selbst an interkulturellen Phänomenen interessiert sein und sich beständig weiterentwickeln wollen. Gemäß der Studie von Pajares (1992) schlagen sich die interkulturellen Ansichten der Lehrperson in ihren Unterrichtspraktiken nieder. Byram (1991) sieht in der interkulturellen Erfahrung der Lehrpersonen die wichtigste Vorbedingung für die erfolgreiche Vermittlung interkultureller Inhalte. Aus der Akkulturationsforschung und aus der Vorurteilsforschung ist jedoch bekannt, dass die interkulturelle Erfahrung zwar eine notwendige, keineswegs aber eine hinreichende Voraussetzung für die Entwicklung interkultureller Kompetenz ist (Berry 1980; Ward 1999; Bennett 1993). Eigene Arbeiten im Rahmen der DESI-Pilotierungsstudie (Göbel 2007) weisen dennoch auf einen deutlichen Zusammenhang zwischen der Auslandserfahrung, der interkulturellen Sensibilität der Lehrkräfte und der realisierten Qualität des interkulturellen Englischunterrichts hin.

Der interkulturelle Lernerfolg von Schülerinnen und Schülern wird durch verschiedene Bedingungen beeinflusst. Hier sind zum einen die allgemeinen Einstellungen der Schülerinnen und Schüler zu interkulturellen Themen zu nennen (Gardener/Lambert 1972; Bennett/Hammer 2002), die einen deutlichen Einfluss auf deren Lernentwicklung nehmen. Auf der anderen Seite stellt die tatsächliche Gelegenheit zur Beschäftigung mit interkulturellen Themen im Unterricht eine wichtige Bedingung für das Lernen dar (siehe auch Carroll 1964; Scheerens/Bosker 1997). Es stellt sich demnach die Frage, inwieweit Schülerinnen und Schüler Gelegenheiten zum interkulturellen Lernen in ihrem Unterricht erhalten und ob sie diese wahrnehmen. Im DESI-Projekt sind die interkulturellen Lernergebnisse der Schülerinnen und Schüler in Anlehnung an das Modell von M. Bennett (1993) durch die Bearbeitung von *Critical Incidents* operationalisiert (siehe Kapitel 17).

In Anlehnung an Überlegungen von Fend (1998), Helmke und Weinert (1997), Helmke (2003) sowie Göbel (2007) wird im Rahmen des DESI-Projektes das in Abbildung 33.1 dargestellte Angebots-Nutzungsmodell zur Modellierung der relevanten Dimensionen zu Grunde gelegt.

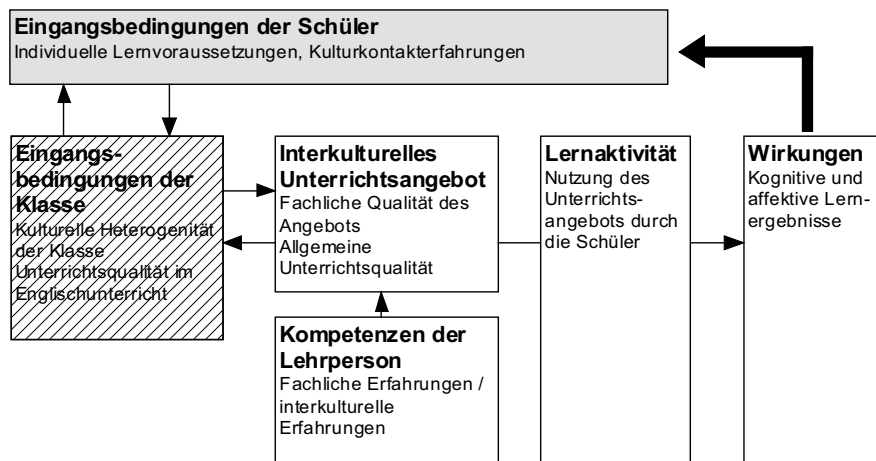


Abbildung 33.1: Angebots-Nutzungs-Modell interkultureller Unterrichtsqualität.

Empirische Untersuchungen zum interkulturellen Englischunterricht, die Aussagen darüber erlauben, wie er in der Schulwirklichkeit aussieht, sind rar. Insofern sind auch deskriptive Aussagen darüber, wie Lehrpersonen und Schülerinnen und Schüler die Bedeutsamkeit interkultureller Themen in ihrem Englischunterricht einschätzen zunächst interessant. Weiterhin interessiert vor dem Hintergrund der theoretischen Vorüberlegungen, welcher Zusammenhang sich zwischen der Bedeutsamkeit interkultureller Themen, den Auslandskontakten, -erfahrungen und den interkulturellen Einstellungen der Lehrpersonen zeigen lässt. Nicht zuletzt soll der Zusammenhang zwischen Lehrermerkmalen, Unterrichtsmerkmalen und den interkulturellen Lernergebnissen der Schülerinnen und Schüler in diesem Kapitel beleuchtet werden.

33.2 Ergebnisse

Angaben der Englischlehrkräfte zur Bedeutsamkeit interkultureller Themen im Englischunterricht

Insgesamt scheinen interkulturelle Themen im Englischunterricht auch in den Augen der in DESI befragten Lehrpersonen sehr wichtig zu sein, im Mittel geben sie an, dass ein prozentualer Anteil von 41% des Unterrichts idealerweise diesem Thema gewidmet werden sollte. Die Angaben der Lehrkräfte variieren zwischen 5-90% der Unterrichtszeit und weisen somit eine deutliche Streubreite und Uneinheitlichkeit in der Einschätzung auf. Der von den Lehrpersonen angegebene tatsächliche Anteil interkultureller Themen im eigenen Unterricht streut zwar in vergleichbarer Weise, liegt im Mittel mit 32% jedoch deutlich darunter. Das Wünschenswerte fällt somit hinter der tatsächlichen Praxis zurück. Um interkulturelle Themen zu behandeln, nutzen ca. drei Viertel der Lehrpersonen selbst entwickelte Materialien (69%). Mehr als

die Hälfte der befragten Lehrpersonen geben an, Video- und Ton-Material für den interkulturellen Englischunterricht zu nutzen (59%). Nur 28% der Lehrpersonen verwenden spezifische Lehrwerke, um das interkulturelle Lernen in ihrem Unterricht zu unterstützen. Hier zeigt sich deutlich die derzeit noch vergleichsweise geringe Verfügbarkeit interkultureller didaktischer Materialien. Bei den verwendeten Lehrwerken handelt es sich hauptsächlich um die im Fachhandel erhältlichen Werke von Cornelsen und Klett. Unabhängig von der Nutzung eigener Materialien wird die Wichtigkeit interkultureller Lernziele von den Lehrpersonen eher hoch eingeschätzt ($M = 2.74$; $SD = 0.26$), was deren Bedeutsamkeit für die Lehrpersonen nochmals unterstreicht.

In Tabelle 33.1 ist der Zusammenhang zwischen der Verwendung unterschiedlicher Lernmaterialien und dem Anteil interkultureller Themen im Unterricht veranschaulicht. Wie zu erwarten, gibt es einen Zusammenhang zwischen der Nutzung bestimmter Unterrichtsmaterialien (spezifische Lehrwerke, eigene Materialien oder Video- und Tonaufnahmen) zur Vermittlung interkultureller Themen und dem prozentualen Anteil den interkulturellen Themen am Unterricht einnehmen. Werden spezifische Lehrmaterialien genutzt, wird auch der Anteil interkultureller Themen am eigenen Unterricht insgesamt etwas höher eingeschätzt. Allerdings ist der Zusammenhang nicht sehr eng und nur für die Nutzung eigener Materialien signifikant ($r = .26$, $p < .01$).

Tabelle 33.1: Prozentualer Anteil interkultureller Themen in Abhängigkeit von der Nutzung spezifischer Unterrichtsmaterialien.

	Nutzung spezifischer interkultureller Lehrwerke ^A		Nutzung eigener Materialien ^A		Nutzung von Video- und Tonaufnahmen ^A	
	Ja	Nein	Ja	Nein	Ja	Nein
Mittlerer Anteil interkultureller Themen	37.6%	26.7%	32.4%	23.5%	33.0%	25.5%

Anmerkungen: ^A ANOVA Gruppenvergleich $p < 0.01$

Hochinferente Beurteilung einer interkulturellen Unterrichtsstunde im Rahmen der DESI-Videostudie

Im Rahmen der DESI-Videostudie erhielten die Lehrpersonen die Vorgabe, eine der beiden aufgezeichneten Stunden einem interkulturellen Thema zu widmen. Diese Vorgabe war ganz allgemein gehalten; den Lehrpersonen wurde es freigestellt, das Thema und die Vorgehensweise selbst zu wählen. Der Unterricht wurde im Hinblick auf die Realisierung des interkulturellen Themas durch mehrere Beurteiler hochinferent auf einer vierstufigen Skala eingeschätzt. Hierbei handelt es sich nur um die Beurteilung der tatsächlichen Realisierung interkultureller Inhalte. Die Realisierung von interkulturellen Inhalten wird dann als gegeben beurteilt, wenn kulturelle und landesspezifische Aspekte eines englischsprachigen Landes (Normen, Werte,

Einstellungen, Sitten, Gebräuche und Verhaltensweisen, Institutionen, geschichtliche Hintergründe) oder damit zusammenhängende geographische Besonderheiten im Unterricht besprochen werden. Die Interraterübereinstimmung bei der Klassifikation der Unterrichtssequenzen ist als sehr gut zu betrachten (Krippendorfs Alpha = .90). Der Mittelwert der Beurteilungen liegt bei 2.81, die Standardabweichung bei 1.33. Mittelwerte und Verteilung der Beurteilungen weisen darauf hin, dass die Lehrpersonen das interkulturelle Thema zum überwiegenden Teil deutlich in dem videografierten Unterricht aufgenommen haben. Entgegen der Vorgabe haben ca. ein Drittel der Lehrpersonen das Thema Interkulturalität überhaupt nicht in der definierten Weise in ihrem Unterricht aufgegriffen. Berichtenswert scheint darüber hinaus, dass der Kulturbezug der interkulturellen Stunde mit dem Kulturbezug der sprachlernorientierten Stunde eine deutliche Korrelation ($r = .32$) aufweist. Dies zeigt, dass die Realisierung der interkulturellen Inhalte in den beiden Stunden nicht unabhängig voneinander ist. Ein Zusammenhang zwischen den Schülereinschätzungen der Bedeutsamkeit und Häufigkeit interkultureller Themen sowie dem Ausmaß der Realisierung interkultureller Themen in der videografierten Unterrichtsstunde zeigte sich hingegen nicht.

Kompetenzen der Lehrkräfte

Interkulturelle Erfahrungen und Auslandskontakte der Englischlehrpersonen

Wie in der Einleitung dargestellt, werden Auslandskontakte von Englischlehrpersonen als bedeutsamer Prädiktor für die Qualität der Realisierung interkultureller Themen im Englischunterricht angesehen (Byram 1991) und sollen hier im Zusammenhang mit dem Unterricht betrachtet werden.

Zu den Auslandserfahrungen der Lehrpersonen ist zunächst zu berichten, dass in diesem Bereich 37% der Lehrpersonen keine Angaben gemacht haben. Wie aber bereits im Kapitel 30 dargestellt, verfügen einige der befragten Lehrpersonen über reichhaltige Erfahrungen im englischsprachigen Ausland. Immerhin geben 90% der Lehrpersonen an, in den letzten zehn Jahren im englischsprachigen Ausland gewesen zu sein. Zumeist handelt es sich um Aufenthalte von bis zu vier Wochen. Darüber hinaus geben 21% der Lehrpersonen an, sieben Monate und länger im englischsprachigen Ausland gewesen zu sein.

Bei der Frage nach aktuellen persönlichen, telefonischen, brieflichen und Email-Kontakten zum englischsprachigen Ausland geben fast ein Drittel der Lehrpersonen an, mehrmals im Jahr persönliche Kontakte zum englischsprachigen Ausland zu haben. Zwischen der Häufigkeit der Auslandsreisen und den Kontakten zum englischsprachigen Ausland zeigt sich erwartungsgemäß ein signifikant positiver Zusammenhang: Je häufiger die Lehrpersonen ins Ausland gereist sind, desto umfangreicher sind die berichteten aktuellen Kontakte ($r = .40$). Weiterhin zeigt sich, dass der angegebene prozentuale Anteil interkultureller Themen am Unterricht signifikant mit den aktuellen Auslandskontakten zusammen hängt ($r = .23$).

Interkulturelle Einstellung der Lehrpersonen

Neben den interkulturellen Erfahrungen der Lehrpersonen werden deren Einstellungen zu interkulturellen Themen für die Realisierung im Unterricht als bedeutsam erachtet (Gardener/Lambert 1972, Bennett/Hammer 2002). Daher soll in Anlehnung an das Modell zur Entwicklung interkultureller Sensibilität von Bennett (1993) die interkulturelle Einstellung der Lehrpersonen dargestellt werden. Die Lehrpersonen haben hierzu einen Multiple-Choice Fragebogen beantwortet, in dem sie ihre subjektive Einschätzung zu interkulturellen Situationen abgeben sollten (AIK-Fragebogen; Allgemeine Interkulturelle Kompetenz). Analog zur Auswertung der Daten der Schülerinnen und Schüler zur Allgemeinen Interkulturellen Sensibilität und zu den interkulturellen Aufgaben im DESI-Projekt (siehe Kapitel 17) wurden die Angaben der Lehrpersonen zur Einschätzung interkultureller Phänomene einer Latenten Klassenanalyse (Latent Class Analysis, LCA) unterzogen. Aus dieser Analyse resultiert ein Profil der interkulturellen Einstellung der Englischlehrkräfte, von dem angenommen wird (Bennett/Bennett/Allen 1999; Göbel 2007), dass es einen Einfluss auf die realisierte interkulturelle Stunde haben könnte. Die aus der LCA resultierende Verteilung der Lehrkräfte auf die unterschiedlichen Klassen interkultureller Einstellungen ist in Abbildung 33.2 dargestellt.

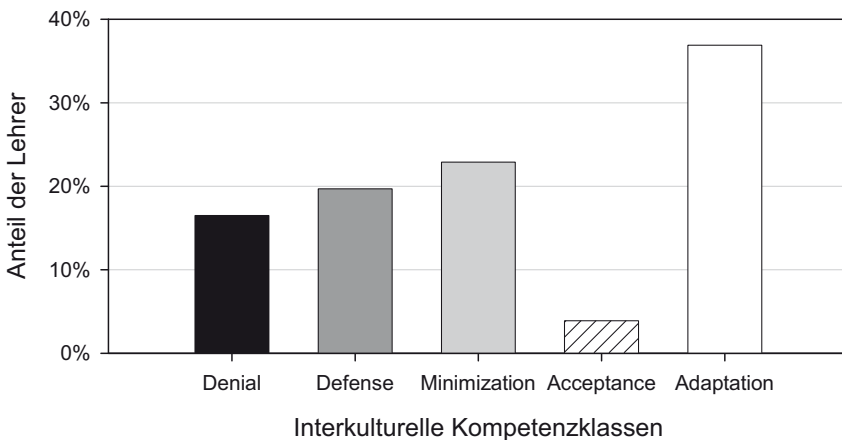


Abbildung 33.2: Zuordnung der Lehrpersonen zu den Typen interkultureller Einstellungen auf Basis der LCA des AIK-Fragebogens (vgl. Hesse/Göbel 2007). Unterschiede in den Gruppengrößen: $\chi^2 = 79.49$; $df=4$; $p < 0.01$, $N = 427$.

Die Lehrpersonen verteilen sich nicht gleichmäßig auf die verschiedenen Gruppen. Nur wenige gehören zur Gruppe Acceptance, demgegenüber gibt es viele Lehrpersonen, die der Adaptation-Gruppe zugeordnet wurden. Im Vergleich zu den Schülern werden weniger Lehrpersonen den Denial- und Defense-Gruppen zugeordnet.

Die Gruppen interkultureller Sensibilität unterscheiden sich vor allem im Hinblick auf die Kontakte zum englischsprachigen Ausland. Hier zeigt sich, dass die Gruppe der Lehrpersonen, die der Denial und Defense Gruppe zugeordnet sind, insgesamt deutlich weniger Kontakte ins englischsprachige Ausland aufweisen als

die anderen Gruppen. Einen prozentual höheren Anteil an interkulturellen Themen in ihrem Unterricht geben die Gruppen Minimization, Acceptance und Adaptation an, diese Unterschiede sind jedoch nicht signifikant. Hinsichtlich der Häufigkeit der Auslandsaufenthalte hat die Defense-Gruppe interessanterweise den höchsten Wert. Tabelle 33.2 gibt einen Überblick über die genannten Gruppenunterschiede.

Tabelle 33.2: Mittelwerte der interkulturellen Sensibilitätsgruppen bezüglich der Skala Auslandskontakte, Anteil interkultureller Themen im Unterricht und Häufigkeit der Auslandsaufenthalte.

	Denial	Defense	Minimization	Acceptance	Adaptation
Auslandskontakte ^A	1.88	2.24	2.34	2.59	2.36
%-Anteil interkultureller Themen	27.09	24.69	34.00	28.00	29.33
Häufigkeit der Auslandsaufenthalte	4.60	6.72	4.66	4.33	5.82

Anmerkungen: ^ADie Mittelwertsunterschiede sind signifikant ($F = 4.12$; $df = 4/212$; $p < .01$).

Einschätzungen der interkulturellen Unterrichtsangebote durch die Schülerinnen und Schüler

Die Häufigkeit, mit der interkultureller Themen im Sinne von Gesprächen über kulturelle Unterschiede in Alltagssituationen im Unterricht vorkommen, wurde auch von Schülerinnen und Schülern auf einer Skala von 1-4 (nie, ein paar Mal pro Jahr, ein Mal pro Monat ein Mal pro Woche) eingeschätzt. Die meisten Lernenden geben hier an, dass interkulturelle Themen mindestens einmal im Monat behandelt werden. Die Schülerinnen und Schüler sollten weiterhin einschätzen, wie wichtig es ihrer Lehrperson ist, dass sie über Verhaltensnormen im englischsprachigen Ausland informiert sind. Hierzu stand eine vierstufige Antwortskala (1 = unwichtig; 4 = sehr wichtig) zur Verfügung. Die Einschätzung der Schülerinnen und Schüler lag im Mittel bei 2.75 ($SD = 0.71$). Sie schätzen das Informiert-sein über Verhaltensnormen im englischsprachigen Ausland also als „eher wichtig“ ein. Allerdings korrelieren die Lehrereinschätzungen und Schülereinschätzungen über das Informiert-sein nicht miteinander ($r = .07$). Gleichmaßen sind der von den Lehrpersonen eingeschätzte prozentuale Anteil interkultureller Themen im Unterricht und die diesbezügliche Häufigkeitseinschätzung durch die Schülerinnen und Schüler unkorreliert ($r = .06$). Zwischen der Wichtigkeit der Themenbehandlung aus Schülersicht und dem Anteil interkultureller Themen aus Lehrersicht gibt es immerhin eine geringe Korrelation ($r = .14$, $p < .05$).

Zusammenhang zwischen Lehrer- und Unterrichtsmerkmalen und den interkulturellen Lernergebnissen

Um zu beurteilen, welchen Einfluss Lehrerkompetenzen und Unterrichtsangebote auf die interkulturellen Lernergebnisse der Schülerinnen und Schüler haben, wurde in mehrbenenanalytischen Modellen der Einfluss ausgewählter Prädiktoren auf das interkulturelle Lernergebnis der Schülerinnen und Schüler berechnet. Kontrolliert wurden hierbei durchweg der Bildungsgang, der individuelle Sprachhintergrund (Erstsprache), der Anteil der Schüler mit nicht deutscher Erstsprache, die kognitiven Grundfähigkeiten der einzelnen Schülerinnen und Schüler sowie der entsprechende Klassendurchschnitt, Geschlecht auf Individualebene sowie der Mädchenanteil in der Klasse, der sozioökonomische Status der Herkunftsfamilie und die soziale Komposition der Klasse. Als Prädiktoren gingen in die Analysen sowohl die für den interkulturellen Englischunterricht spezifischen Dimensionen der Unterrichtsqualität als auch allgemeine Dimensionen der Unterrichtsqualität wie Klassenführung und Fehlerkultur ein. Klassenführung bildet die Nutzung der Zeit für unterrichtsrelevante Themen ab und wird in verschiedenen Untersuchungen (z.B. Wang/Haertel/Walberg 1993; Kounin 1973) immer wieder als für den Lernprozess relevante Einflussgröße beschrieben. Eine positive Fehlerkultur (Oser/Hascher 1997) gilt als positive klimatische Bedingung des Unterrichts einerseits und als Anregungsbedingung für die Beförderung von komplexen Lernprozessen andererseits und kann insofern einen allgemein förderlichen Einfluss auf die Lernergebnisse der Schüler haben. Die interkulturellen Lernergebnisse werden durch die Ergebnisse der Bearbeitung der interkulturellen Aufgaben (*Critical Incidents*) durch die Schülerinnen und Schüler operationalisiert (Kapitel 17, sowie Hesse/Göbel 2007). In den Abbildungen 33.3 und 33.4 sind die Zusammenhänge zwischen Prädiktoren und Lernergebnis der Analysen grafisch veranschaulicht: Auf der x-Achse sind die Prädiktoren auf Klassenebene abgetragen, auf der y-Achse die Zuordnungswahrscheinlichkeiten der Schüler zu den jeweiligen interkulturellen Kompetenzklassen (auf der Grundlage der Bearbeitung der interkulturellen Aufgaben – *Critical Incidents*). Für die einzelnen Prädiktoren wurden separate Analysen durchgeführt. Für die schwach gedruckten Prädiktoren zeigen sich keine signifikanten Unterschiede in der Zuordnungswahrscheinlichkeit zu den interkulturellen Kompetenzklassen.

Insgesamt zeigen die Analyseergebnisse, dass der Kontakt der Lehrperson mit dem englischsprachigen Ausland einen schwachen, aber signifikanten Einfluss auf die interkulturellen Lernergebnisse der Lernenden hat. Der Unterschied in den Zuordnungswahrscheinlichkeiten zwischen den interkulturellen Kompetenzklassen ist zwar nicht sehr groß, aber er ist zwischen den Acceptance/Adaptation Schülern und den Denial/Defense Schülern signifikant verschieden. Wenn eine Lehrperson mehr Auslandskontakte hat, dann erhöht sich für die Schüler der Klasse die Wahrscheinlichkeit, bei Bearbeitung der interkulturellen Aufgaben (*Critical Incidents*) zur Acceptance/Adaptation-Gruppe zugeordnet zu werden. Neben den interkulturellen Erfahrungen der Lehrperson hat ein positiver Umgang mit Fehlern einen günstigen Einfluss auf die interkulturellen Lernergebnisse. Wird von den Schülern wahr-

genommen, dass Fehler im Unterricht von der Lehrperson als Lerngelegenheiten genutzt werden und ein Klima herrscht, in dem Fehler erlaubt sind, ja gar produktiv genutzt werden, dann zeigen auch hier die Schülerinnen und Schüler bessere interkulturelle Lernergebnisse.

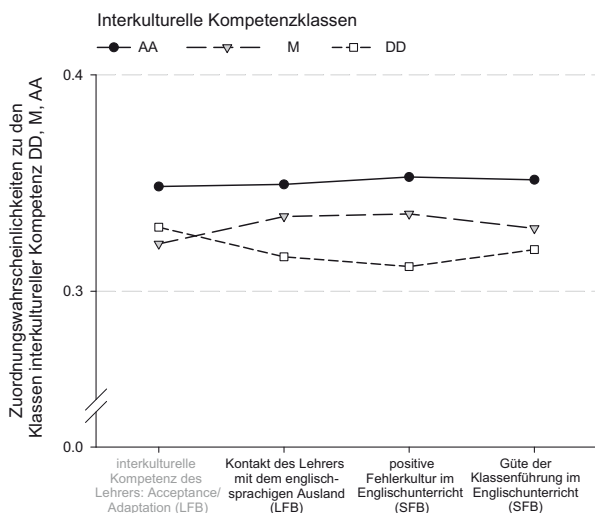


Abbildung 33.3: Zuordnungswahrscheinlichkeiten zu den Typen interkultureller Einstellungen in Abhängigkeit der Unterrichtsmerkmale aus Schüler- und Lehrpersonen-sicht. Merkmale ohne signifikanten Effekt sind in der Beschriftung der Abszisse grau gedruckt.

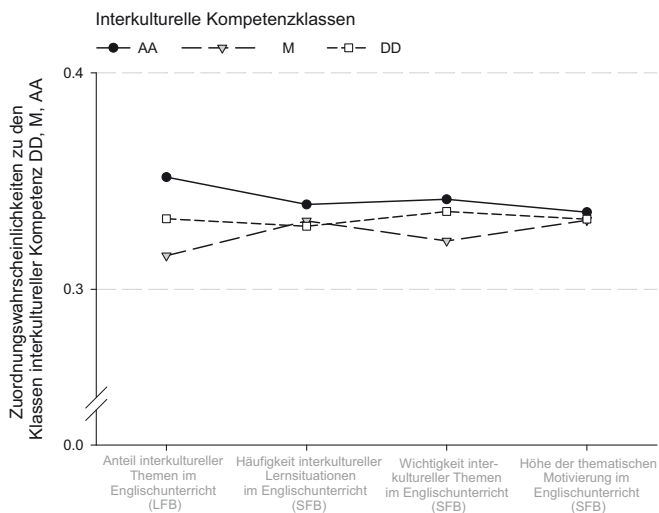


Abbildung 33.4: Zuordnungswahrscheinlichkeiten zu den Typen interkultureller Einstellungen in Abhängigkeit der Unterrichtsmerkmale aus Schüler- und Lehrpersonen-sicht. Merkmale ohne signifikanten Effekt sind in der Beschriftung der Abszisse grau gedruckt.

Darüber hinaus hat eine gute Klassenführung im Sinne der aktiven Lernzeitnutzung ebenfalls einen signifikant positiven Einfluss auf das interkulturelle Lernergebnis. Demgegenüber zeigt sich jedoch weder beim Anteil interkultureller Themen am Gesamtunterricht, noch bei der von den Schülern eingeschätzten Wichtigkeit interkultureller Themen im Englischunterricht ein signifikanter Einfluss auf die interkulturellen Lernergebnisse. Weiterhin zeigt sich bei der Wahrnehmung der themenspezifischen Motivierung kein signifikanter Einfluss auf die interkulturellen Lernergebnisse der Schüler (vgl. Abbildungen 33.3 und 33.4). Auch die interkulturelle Einstellung der Lehrperson schlägt sich kaum auf die Lernergebnisse der Schülerinnen und Schüler nieder. Bei der Interpretation der Ergebnisse muss jedoch berücksichtigt werden, dass die Angebote zum Teil aus Schülersicht beurteilt wurden. Bereits die geringen Korrelationen zwischen Lehrer- und Schülerurteil im Hinblick auf Wichtigkeit und Lerngelegenheiten von interkulturellen Inhalten weisen darauf hin, dass die Lernenden nicht immer die Absichten ihrer Lehrpersonen erkennen.

33.4 Diskussion

Die deskriptiven Befunde zeigen zunächst, dass interkulturelle Themen im Englischunterricht von den meisten Lehrpersonen als wichtig eingeschätzt werden und auch einen beträchtlichen Teil der Unterrichtszeit einnehmen (im Mittel 35%). In gleicher Weise zeigt sich im Rahmen der videografierten Englischstunde, die einem interkulturellen Thema gewidmet sein sollte, dass ca. 60% der Lehrpersonen dieses Thema deutlich realisieren. Dies kann als Bestätigung dafür gewertet werden, dass interkulturelle Themen im Englischunterricht tatsächlich realisiert werden können, wenn entsprechende Vorgaben gemacht werden. Darüber hinaus nehmen einige Lehrpersonen interkulturelle Aspekte auch in ihren sprachlernorientierten Stunden auf. Auch Schülerinnen und Schüler erachten interkulturelle Themen im Englischunterricht als bedeutsam, wenngleich ihre Einschätzung mit der ihrer Lehrpersonen nur wenig Übereinstimmung aufweist. Die geringen Korrelationen zwischen Schüler- und Lehrpersoneneinschätzung der Wichtigkeit von interkulturellen Inhalten und der Lerngelegenheiten weisen darauf hin, dass die Schüler nicht immer die Intentionen ihrer Lehrpersonen erkennen können.

Ein Zusammenhang zwischen den Auslandserfahrungen der Lehrperson und interkulturellem Englischunterricht lässt sich zunächst insofern herstellen, als die Lehrpersonen, die über mehr aktuelle Kontakte ins englischsprachige Ausland verfügen, interkulturellen Themen in ihrem Unterricht mehr Zeit einräumen. Lehrpersonen, die im Hinblick auf ihre interkulturelle Einstellung der Denial-Gruppe zugeordnet sind, widmen interkulturellen Themen in ihrem Unterricht weniger Zeit als alle anderen Gruppen. Zwischen den realisierten Themen in der videografierten Stunde und den Einstellungen und Erfahrungen der Lehrpersonen zeigte sich nur ein schwacher Zusammenhang, Lehrpersonen der Denial-Gruppe realisierten im Mittel etwas weniger deutlich interkulturelle Themen im videografierten Unterricht.

Im Hinblick auf die interkulturellen Lernergebnisse hat sich die interkulturelle Erfahrung der Lehrkräfte als wichtig herausgestellt. An dieser Stelle bestätigen sich Ergebnisse aus der DESI-Pilotierung (vgl. Göbel 2007). Die Erfahrungsbasiertheit interkulturellen Englischunterrichts wurde als ein wichtiges Qualitätsmerkmal herausgearbeitet, das sich positiv auf interkulturelle Lernergebnisse auswirken kann. Je mehr Auslandskontakte Lehrkräfte haben, desto mehr Raum können sie den Schülern geben, eigene Erfahrungen einzubringen und zu reflektieren. Dies ist eine wichtige Unterrichtsqualitätsdimension im interkulturellen Englischunterricht. Darüber hinaus haben sich auch allgemeine didaktische Kompetenzen wie eine positive Fehlerkultur und ein gutes Classroom-Management als bedeutsam für die interkulturellen Lernergebnisse erwiesen. Es kommt also auch im interkulturellen Englischunterricht sowohl auf allgemeine didaktische Kompetenzen als auch auf interkulturelle Erfahrungen an, um die Schülerinnen und Schüler in ihren Lernprozessen anzuregen. Die geringe Bedeutsamkeit anderer unterrichtlicher Variablen für die interkulturelle Kompetenz legt die Vermutung nahe, dass neben den individuellen Merkmalen wie den kognitiven Fähigkeiten oder der Erstsprache weitere außerschulische Einflussfaktoren wirksam sind. Bei der Interpretation der Ergebnisse muss jedoch berücksichtigt werden, dass die fachlichen Angebote zum Teil aus Schülersicht beurteilt werden. Hieran zeigt sich wiederum, dass Schüler die fachliche Qualität des Unterrichts weniger gut einschätzen können, als allgemeine didaktische Kompetenzen der Lehrkraft.

Die Auswertung der DESI-Videos zur Vermittlung interkultureller Inhalte im Englischunterricht ist noch nicht abgeschlossen. Die Auswertung der interkulturellen Qualität der videografierten Unterrichtsstunden der Hauptuntersuchung erfolgte bislang nur in sehr grober Weise. Es wird in weiteren Auswertungen darum gehen müssen, die interkulturelle Qualität des Unterrichts differenzierter zu bestimmen, um dadurch einerseits den Einfluss der interkulturellen Erfahrungen der Lehrkräfte differenzierter auf den Unterricht beziehen zu können und andererseits den Einfluss des Unterrichts auf die Lernergebnisse der Schülerinnen und Schüler in den interkulturellen Aufgaben darstellen zu können. Die Analyse der Praxis unterschiedlicher Vermittlungsansätze kann dann auch als eine hervorragende Grundlage angesehen werden, um Entwicklungen im Fach Englisch voranzutreiben und damit eine systematische Konzeptionierung der interkulturellen Unterrichtsinhalte und Materialien zu ermöglichen (siehe auch Met/Galloway 1992).

Literatur

- Baron, R. S. (2002): *Interculturally Speaking*. München: Langenscheidt ELT.
- Bennett, J./Bennett, M./Allen, W. (1999): Developing Intercultural Competence in the Language Classroom. In: Paige, R. M./Lange, D. L./Yershova, Y. A. (Hrsg.): *Culture as the Core. Integrating Culture into the Language Curriculum*. Minneapolis: Center for Advanced Research on Language Acquisition, S. 13-45.

- Bennett, M. J. (1993): Towards Ethnorelativism: A Developmental Model of Intercultural Sensitivity. In: Paige, M. (Hrsg.): Education for the intercultural experience. Yarmouth: Intercultural Press, S. 21-72
- Bennett, M. J./Hammer, M. R. (2002): The Intercultural Development Inventory (IDI). Portland: International Communication Institute.
- Berry, J. W. (1980): Social and cultural change. In: Triandis, H. /Brislin, R. W. (Hrsg.): Handbook of Cross-Cultural Psychology. Boston: Allyn and Bacon, Vol. 5, S. 211-279.
- Burwitz-Meltzer, E. (2003): Allmähliche Annäherungen: Fiktionale Texte im interkulturellen Fremdsprachenunterricht der Sekundarstufe I. Tübingen: Günter Narr Verlag.
- Byram, M. (1991): Teaching culture and language: towards an integrated model. In: Buttjes, D./Byram, M. (Hrsg.): Mediating languages and cultures: Towards an intercultural theory of foreign language education. Clevedon, Avon, England: Multilingual Matters, S. 17-30.
- Carroll, J. B. (1964): Research on Teaching Foreign Languages. In: Gage, N. L. (Hrsg.): Handbook of Research on teaching, Vol. 21. Chicago: Rand McNally, S. 1060-1099.
- Damen, L. (1987): Culture Learning. The Fifth Dimension in the Language Classroom. Reading: Addison-Wesley.
- Edelhoff, Ch. (1987): Lehrerfortbildung und interkulturelles Lehren und Lernen. In: Baumgartz, G. / Stephan, R. (Hrsg.): Fremdsprachenlernen als Beitrag zur internationalen Verständigung. Inhaltliche und organisatorische Perspektiven der Lehrerfortbildung in Europa. München: Judicium, S. 110-148.
- Fend, H. (1998): Qualität im Bildungswesen. Weinheim: Juventa.
- Gardner, R. C./Lambert, W. E. (1972): Attitudes and motivation in second language learning. The role of attitudes and motivation. London: Edward Arnold.
- Göbel, K./Hesse, H. G. (2004): Vermittlung interkultureller Kompetenz im Englischunterricht – eine curriculare Perspektive. In: Zeitschrift für Pädagogik 50, S. 818-834.
- Göbel, K. (2007): Unterrichtsqualität im interkulturellen Englischunterricht – eine Videostudie. Münster: Waxmann.
- Helmke, A./Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In Weinert, F. E. (Hrsg.): Enzyklopädie der Psychologie. Serie Pädagogische Psychologie, Band 3: Psychologie des Unterrichts und der Schule. Göttingen: Hogrefe, S. 71-176.
- Helmke, A. (2003): Unterrichtsqualität: Erfassung, Bewertung, Verbesserung. Velber: Kallmeyersche Verlangsbuchhandlung.
- Hesse, H.-G./Göbel, K. (2007): Interkulturelle Kompetenz. In: Beck, B./Klieme, E. (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band I. Weinheim: Beltz, S. 256-272.
- Kultusministerkonferenz (2003): Bildungsstandards für die erste Fremdsprache (Englisch/ Französisch) für den Mittleren Schulabschluss. Beschluss der Kultusministerkonferenz vom 4.12.2003.
- Kounin, J. S. (1976): Techniken der Klassenführung. Bern: Hans Huber Verlag.
- Met, M./Galloway, V. (1992): Research in foreign language curriculum. In: Jackson, P. W. (Hrsg.): Handbook of Research on Curriculum. New York: Macmillan, S. 852-890.
- Oser, F./Hascher, T. (1997): Lernen aus Fehlern. Schriftenreihe zum Projekt „Lernen Menschen aus Fehlern? Zur Entwicklung einer Fehlerkultur in der Schule“, Nr. 1. Freiburg: Pädagogisches Institut.
- Paige, M. R./Jorstad, H./Siaya, L./Klein, F./Colby, J. (1999): Culture learning in language education: a review of the literature. In: Culture as the core: integrating culture into the language curriculum. Minneapolis, MN: Center for advanced research on language acquisition. CARLA working Paper No.15, S. 47-114.
- Pajares, F. M. (1992): Teacher's beliefs and educational research: Cleaning up a messy construct. In: Review of Educational Research 62, S. 307-322.

- Scheerens, J./Bosker, R. J. (1997): *The Foundations of Educational Effectiveness*. Oxford: Elsevier Science.
- Smith, S. L./Paige, M./Steglitz, I. (2003): *Theoretical Foundations of Intercultural Training and Applications to the Teaching of Culture*. In: Lange, D. L. /Paige, M. (Hrsg.): *Culture as the Core*. Greenwich: IAP, S. 89-126.
- Steinert, B./Klieme, E. (2004): Was kommt mit der Einführung der Bildungsstandards auf die Schulen zu? In: *Schulverwaltung NI SH 2*, S. 36-39.
- Trim, J./North, B./Coste, D./Sheils, J. (Hrsg.) (2001): **Gemeinsamer Europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen**. Berlin: Langenscheidt.
- Vollmer, H. J. (2001): *Wissenschaftlich fundierte Kompetenzmodelle im Bereich der Sprachen und deren Niveaudifferenzierung / Skalierung (am Beispiel des „Referenzrahmens“ des Europarats)*. Unpubliziertes Manuskript: Osnabrück.
- Wang, C. M./Haertel, G. D./Walberg, H. (1993): *Toward a Knowledge Base for School Learning*. In: *Review of Educational Research* 63, S. 249-294.
- Ward, C./Rana-Deuba, A. (1999): *Acculturation and Adaptation Revisited*. In: *Journal of Cross-Cultural Psychology* 30, S. 422-439.

Brigitte Steinert / Johannes Hartig / Eckhard Klieme

34 Institutionelle Bedingungen der Sprachkompetenzen

34.1 Sprachliche Kompetenzen im Deutschen und Englischen im Schulartvergleich

Über die Unterschiede in den sprachlichen Kompetenzen im Deutschen und im Englischen zwischen den Bildungsgängen, denen die Schülerinnen und Schüler individuell zugeordnet sind, wurde bereits in den Kapiteln 5 bis 16 berichtet. Erwartungsgemäß lagen in allen in DESI erfassten Kompetenzbereichen die Leistungen im gymnasialen Bildungsgang über jenen im mittleren Bildungsgang (Realschule), die wiederum die Leistungen im Hauptschul-Bildungsgang und in den Integrierten Gesamtschulen übertrafen. Für die Analyse der institutionellen Bedingungen der Sprachkompetenzen werden in diesem Kapitel zunächst die Mittelwerte und Streuungen der Schülerkompetenzen in den einzelnen Schularten ermittelt, um die Unterschiede und Überschneidungen in den Leistungsverteilung im Deutschen und im Englischen zwischen den Schularten und im Vergleich der Fächer aufzuzeigen. Die Fokussierung auf Schularten anstelle von Bildungsgängen hat ihre Begründung darin, dass die Schularten die organisatorischen Einheiten sind, auf die sich institutionelle Zielsetzungen, Regelungen und Praktiken von Schulleitungen und Fachkollegien beziehen. Aus diesem Grund werden die Schulen mit mehreren Bildungsgängen, die den Hauptschul- und Realschulbildungsgang innerhalb einer Schule kombinieren, gesondert aufgeführt.

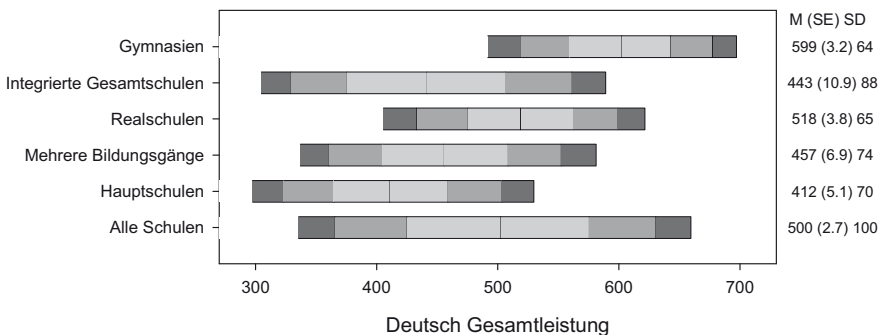


Abbildung 34.1: Perzentilbänder der Deutschkompetenz nach Schularten.

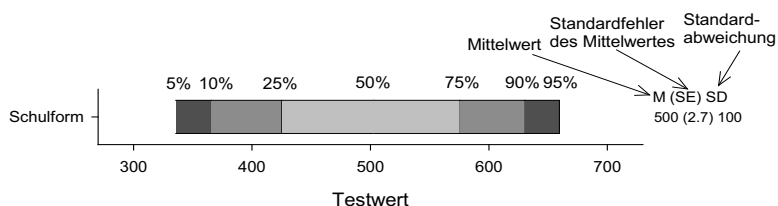
Die Verteilungen der sprachlichen Kompetenzen im Deutschen und im Englischen in den einzelnen Schularten sind in Abbildung 34.1 und Abbildung 34.2 an-

hand von Perzentilwerten und Perzentilbändern¹ dargestellt. Der Mittelwert der Gesamtleistung aller Schülerinnen und Schüler ist auf 500 Punkte normiert worden, die Standardabweichung auf 100 Punkte. Die Schularten, in denen die Schülerkompetenzen im Deutschen signifikant *über* dem Gesamtdurchschnitt von 500 Punkten liegen, sind die Realschulen mit 518 Punkten und die Gymnasien mit 599 Punkten. Die Schularten, in denen die Schülerkompetenzen im Deutschen signifikant *unter* dem Gesamtdurchschnitt von 500 Punkten liegen, sind die Schulen mit mehreren Bildungsgängen mit 457 Punkten, die Integrierten Gesamtschulen mit 443 Punkten und die Hauptschulen mit 412 Punkten.

Die Streuung der Deutschkompetenzen ist in den Gesamtschulen und den Schulen mit mehreren Bildungsgängen erwartungsgemäß am höchsten. Die Leistungsunterschiede zwischen diesen beiden Schularten sind bei den Deutschkompetenzen nicht signifikant. Gegenüber den Hauptschulen weisen beide Schularten einen signifikanten Vorsprung auf, auch wenn dieser nicht allzu groß ist. Nach Abbildung 34.1 würden im Fach Deutsch die 10% leistungsstärksten Hauptschüler in der Realschule zur Leistungsmitte gehören und im Gymnasium noch im untersten Quartil anzufinden sein. Die 10% leistungsstärksten Schülerinnen und Schüler in den Schulen mit mehreren Bildungsgängen und den Gesamtschulen gehörten zur unteren Leistungsmitte im Gymnasium, die 10% leistungsstärksten Realschüler zur oberen Leistungsmitte im Gymnasium.

Die Verteilung der Schülerleistungen im Englischen zeigt im Vergleich zu den Deutschleistungen zum Teil ein anderes Muster (vgl. Abbildung 34.2). Signifikant *über* dem Gesamtdurchschnitt von 500 Punkten liegen auch hier die Englischleistungen in den Realschulen mit 518 Punkten und in den Gymnasien mit 611 Punkten. Signifikant *unter* dem Gesamtdurchschnitt von 500 Punkten liegen die Englischkompetenzen in den Integrierten Gesamtschulen mit 438 Punkten, den Schulen mit mehreren Bildungsgängen mit 424 Punkten sowie in den Hauptschulen mit 417 Punkten.

- 1 Die Verteilungen der Schülerleistungen werden anhand von Perzentilbändern dargestellt. Hierbei werden die Spannweite und die Verteilung der Leistungswerte veranschaulicht. Diese geben an, wie viel Prozent der Schüler in bestimmten Leistungsbereichen liegen. So bedeutet das 5. Perzentil, dass 5% der Schüler noch unter diesem Wert liegen, während das 50. Perzentil den Wert angibt, der die Verteilung in zwei gleich große Teile teilt. Zwischen dem 25. und 75. Perzentil liegen die hinsichtlich ihrer Leistungen mittleren 50% der Schüler. Zusätzlich sind rechts am Rand der Abbildungen die Mittelwerte, deren Standardfehler und die Standardabweichungen der Leistungen in der jeweiligen Schulart angegeben. Die folgende Abbildung gibt eine Lesehilfe für die Perzentilband-Darstellungen:



Am *unteren Ende der Leistungsskala* sind die Verteilungen im Englischen gestaucht: zwischen dem fünften und dem fünfundzwanzigsten Perzentil liegen im Deutschen bei Hauptschulen, Schulen mit mittleren Bildungsgängen und Integrierten Gesamtschulen jeweils ca. 70 Testpunkte, im Englischen jedoch nur ca. 45 Punkte. In diesen drei Schularten gibt es eine gewichtige Gruppe von Schülerinnen und Schülern, deren Sprachkompetenzen im Englischen deutlich hinter den curricula- ren Erwartungen zurückbleiben und die bei den schriftlichen DESI-Tests minimale Leistungen erbringen.

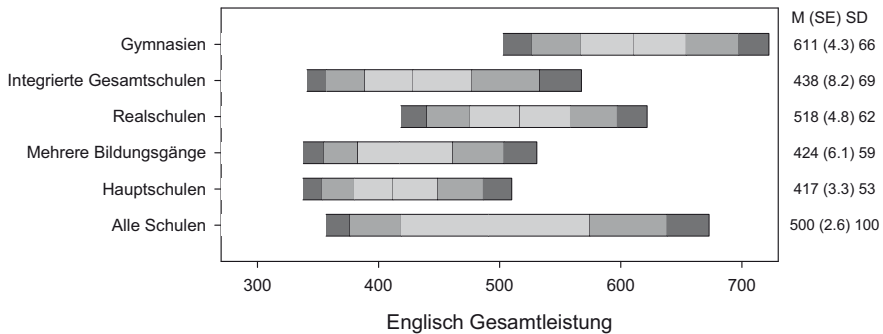


Abbildung 34.2: Perzentilbänder für die Englischkompetenz nach Schularten.

Die *Unterschiede zwischen den Schularten* sind im Englischen größer als im Deutschen. Bei den Deutschkompetenzen würden die 10% leistungstärksten Schülerinnen und Schüler der Hauptschulen in den Realschulen zur Leistungsmitte gehören und auch in den Gymnasien noch im untersten Quartil anzufinden sein. Bei den Englischkompetenzen gilt dies nicht: hier überschneiden sich die Perzentilbänder der Hauptschulen und Gymnasien kaum noch. Die *Schulen mit mehreren Bildungsgängen* befinden sich hinsichtlich der Schülerleistungen im Deutschen genau zwischen Hauptschulen und Realschulen. Im Englischen unterscheidet sich ihr Leistungsdurchschnitt jedoch nicht signifikant von dem der Hauptschulen und liegt deutlich unter Realschulniveau. Offenbar ist der Englischunterricht in diesen Schulen weniger gut entwickelt als der Deutschunterricht. Die *Gesamtschulen* weisen im Deutschen wie auch im Englischen erwartungsgemäß die höchste Leistungsstreuung auf. In den mittleren Deutschleistungen unterscheiden sich die Integrierten Gesamtschulen nicht signifikant von den Hauptschulen. Im Englischen weisen sie jedoch gegenüber den Hauptschulen einen kleinen, aber signifikanten Vorsprung auf. Die 10% leistungstärksten Schülerinnen und Schüler der Gesamtschulen würden zur oberen Leistungsmitte in den Realschulen und noch zum untersten Quartil in den Gymnasien gehören. In den *Gymnasien* gibt es im Englischen eine Leistungsspitze von etwa 10% der Schülerinnen und Schüler, die mehr als zwei Standardabweichungen über dem Durchschnitt der Gesamtpopulation liegt. Bei einer insgesamt sehr hohen Leistungsstreuung zwischen den Gymnasien setzt sich diese Leistungsspitze mit einem deutlich größeren Vorsprung von den übrigen Schularten ab als die Leistungsspitze im Deutschen.

Der Vergleich der Mittelwerte und Streuungen der Schülerleistungen im Deutschen und im Englischen zeigt die institutionellen Unterschiede in den Sprachkompetenzen zwischen den Schularten anschaulich auf. Um das Ausmaß der einzelschulischen Gestaltungsspielräume abzuschätzen, stellt sich die Frage, wie groß die institutionelle Variabilität zwischen den Schularten, zwischen den Schulen und innerhalb der einzelnen Schularten jeweils ist.

34.2 Schulunterschiede in den sprachlichen Kompetenzen

Im Rahmen der PISA-Studien hat sich gezeigt, dass in Staaten mit einem intern differenzierten Sekundarbereich ein größerer Anteil der Gesamtvariation in den Schülerkompetenzen auf Unterschiede innerhalb von Schulen zurück geht als in Staaten mit gegliedertem Sekundarbereich wie Deutschland, die eine externe Leistungsdifferenzierung im Schulsystem vornehmen (vgl. Baumert/Schümer 2001; Prenzel/Senkbeil/Drechsel 2004). Ferner gibt es Hinweise darauf, dass die schulart-spezifischen Unterschiede in den Schülerleistungen im Laufe der Sekundarschulzeit zunehmen und die Unterschiede zwischen Schulen derselben Schulart abnehmen, so dass die durchschnittliche Leistungsunterschiede zwischen Schulen am Ende der Sekundarstufe I eher moderat ausfallen (Baumert/Trautwein/Artelt 2003). In DESI ist ein internationaler Vergleich der Leistungsvariation im Deutschen und im Englischen nicht möglich. Die Stichprobe erlaubt aber eine Antwort auf die Frage, wie groß die Variabilität der Schülerkompetenzen im sprachlichen Bereich auf zentralen Ebenen des Schulsystems ist. Mit der Zerlegung der Gesamtvariation der Sprachkompetenzen in individuelle und institutionelle Varianzkomponenten kann Auskunft darüber gegeben werden, in welchem Ausmaß die Gesamtvariation der Schülerkompetenzen im Deutschen und im Englischen auf Unterschiede *innerhalb von Schulen*, auf Unterschiede *zwischen den Schulen* und auf Unterschiede *zwischen den Schularten* zurückgeführt werden kann. Alle Schätzungen der Varianzkomponenten erfolgten in Mehrebenenmodellen mit HLM 6 (Raudenbusch/Bryk/Congdon 2005) und auf der Basis der Restricted Maximum Likelihood-Schätzung.

Bei den Sprachkompetenzen im Deutschen geht mehr als die Hälfte der Gesamtvarianz auf Unterschiede zwischen den Schularten zurück (vgl. Abbildung 34.3). 11% der Gesamtvarianz entfallen auf die Kompetenzunterschiede zwischen den Schulen derselben Schulart, 37% auf Unterschiede innerhalb der Schulen derselben Schulart. Damit sind insgesamt knapp zwei Drittel der Gesamtvariation der Schülerkompetenzen im Deutschen an institutionelle Komponenten gebunden und gut ein Drittel an individuelle. Auch wenn die durchschnittliche Zwischenschulvarianz von 11% eine relative Homogenität der leistungsmäßigen Zusammensetzung der Schülerschaft in Schulen derselben Schulart indiziert, sind diese Unterschiede durchaus bedeutsam.

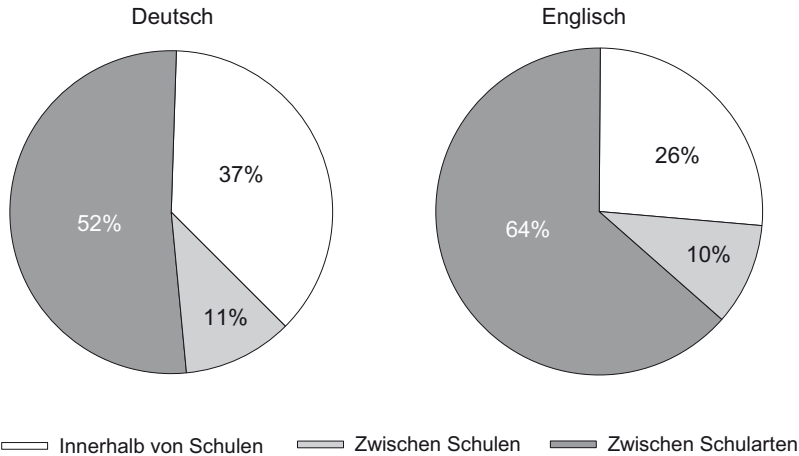


Abbildung 34.3: Individuelle und institutionelle Varianzkomponenten der Sprachkompetenzen im Deutschen und im Englischen.

Im Englischen gehen mit 64% knapp zwei Drittel der Gesamtvarianz der sprachlichen Kompetenzen auf Unterschiede zwischen den Schularten zurück (vgl. Abbildung 34.3). 10% der Gesamtvarianz beruhen auf Kompetenzunterschieden zwischen Schulen derselben Schulart, 26% auf Unterschieden innerhalb der Schulen derselben Schulart. Damit sind im Englischen fast drei Viertel der gesamten Kompetenzstreuung an institutionelle Komponenten gebunden und nur ein Viertel an individuelle. Da die Kompetenzen im Englischen in höherem Maße im schulischen Lernkontext erworben werden als die Sprachkompetenzen im Deutschen, das für die meisten Schülerinnen und Schüler in der DESI-Stichprobe die Erstsprache ist, sind die durch die Schulartzugehörigkeit erklärten Kompetenzunterschiede erwartungsgemäß größer.

Um einen Anhaltspunkt für die Einschätzung der institutionellen Variabilität der Schülerkompetenzen im Deutschen und im Englischen zu erhalten, sind in Tabelle 34.1 die entsprechenden Varianzanteile für die Lese- und Mathematikkompetenz der Schülerinnen und Schüler der neunten Jahrgangsstufe aus PISA 2000-E als Vergleichswerte aufgeführt (vgl. Baumert/Trautwein/Artelt 2003). Aus dieser Übersicht geht hervor, dass die Kompetenzunterschiede im Deutschen und im Englischen in DESI in höherem Maße auf Schulartunterschiede (52% und 64%) zurückzuführen sind als die Unterschiede in der Lese- und mathematischen Kompetenz (45% und 47%) in PISA 2000. Das größere Ausmaß der Schulartunterschiede in DESI dürfte eine Folge der curricularen Orientierung der DESI-Testaufgaben sein, während die Testaufgaben zur Lesekompetenz und zur Mathematik in PISA stärker funktional am Literacy- bzw. Grundbildungskonzept orientiert sind (Deutsches PISA Konsortium 2001, 2003). Deswegen zeigt sich in PISA eine größere Leistungsheterogenität innerhalb der Schulen. Hinsichtlich der Variabilität, die nur auf Unterschiede zwischen den Einzelschulen zurückgeführt kann, liegen die entsprechenden Varianzanteile in DESI mit 11% für Deutsch und 10% für Englisch sehr nahe bei den Werten für

die Lesekompetenz in PISA (12%). Der niedrigere Varianzanteil von 7% für die Unterschiede zwischen den Einzelschulen in der Mathematik könnte ein Hinweis darauf sein, dass der Mathematikunterricht innerhalb der Schularten stärker vereinheitlicht ist.

Tabelle 34.1: Individuelle und institutionelle Varianzkomponenten der Kompetenzen von Neuntklässlern nach Kompetenzbereichen in DESI und PISA 2000-E (in % der Gesamtvarianz).

Stichprobe	Kompetenzbereich	Innerhalb von Schulen	Zwischen Schulen	Zwischen Schularten
DESI	Deutsch	37	11	52
DESI	Englisch	26	10	64
PISA 2000-E	Lesekompetenz	43	12	45
PISA 2000-E	Mathematik	46	7	47

Eine durchschnittliche Zwischenschulvarianz der Schülerkompetenzen im Deutschen und im Englischen von rund 10% als Maß für die Variabilität zwischen Einzelschulen deutet auf ein relativ geringes Potential für die einzelschulische Qualitätsentwicklung gegenüber individuellen und schulartbezogenen Gestaltungsparametern hin. Gleichwohl sind die Kompetenzunterschiede zwischen den Einzelschulen nicht unbedeutend. Wenn die Varianzkomponenten der sprachlichen Kompetenzen innerhalb jeder einzelnen Schulart ermittelt werden, ist die Streuung in den Schülerkompetenzen innerhalb der einzelnen Schularten notwendigerweise kleiner als die Streuung zwischen den Schulen aller Schularten. Die Variabilität der Schülerkompetenzen zwischen den einzelnen Schulen innerhalb der Schularten gibt gleichwohl einen Hinweis darauf, in welchem Ausmaß die Schülerkompetenzen innerhalb jeder Schulart einzelschulischen Gestaltungspraktiken zugänglich sein könnten.

Aus Abbildung 34.4 geht hervor, dass die Gymnasien in Bezug auf die Streuung der Schülerkompetenzen im Deutschen, die mit einem Varianzanteil von 14% auf Unterschiede zwischen Schulen zurückgeht, die Schulart mit der am stärksten ausgeprägten Homogenität sind und die Integrierten Gesamtschulen, die strukturell eine leistungsheterogene Schülerschaft bedienen, mit einer Zwischenschulvarianz von 38% die Schulart mit der am stärksten ausgeprägten Heterogenität. Auch in den Hauptschulen ist die institutionelle Variabilität der Deutschkompetenzen mit einem Varianzanteil von 23% recht hoch.

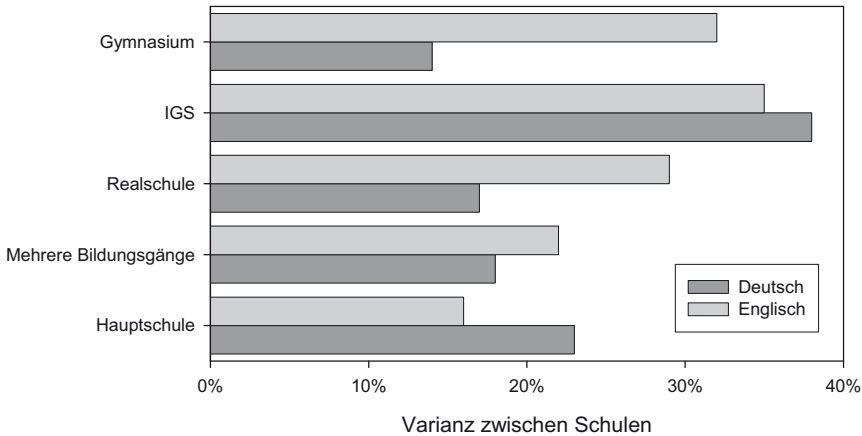


Abbildung 34.4: Varianzanteile der Schülerkompetenzen im Deutschen und im Englischen zwischen Schulen innerhalb der einzelnen Schularten (in %).

Im Englischen zeigt sich das umgekehrte Bild. Hier weisen die Gymnasien – nach den Integrierten Gesamtschulen – die größten Unterschiede zwischen den Einzelschulen auf und die Hauptschulen die geringsten. 32% der Variation der Englischkompetenzen gehen im Gymnasium auf Unterschiede zwischen den einzelnen Schulen zurück, in der Hauptschule nur 16%. Diese Unterschiede in der institutionellen Variabilität der Englischleistungen veranschaulichen die in den Perzentilband-Darstellungen gefundenen Verteilungsmuster: die niedrigere Streuung der Englischkompetenzen in den Hauptschulen und den Schulen mit mehreren Bildungsgängen wird um den Preis insgesamt unterdurchschnittlicher Schülerleistungen erzielt, während die insgesamt überdurchschnittlichen Englischleistungen in den Gymnasien mit einer großen Leistungsspreizung einhergehen. Im Deutschen erzielen die Gymnasien überdurchschnittliche Schülerleistungen in weitaus homogeneren Lernkontexten. Inwieweit die fachspezifischen Differenzen in der Leistungsstreuung zwischen Schulen derselben Schulart mit Unterschieden in der Konzeption und Standardisierung der Curricula oder mit Unterschieden zwischen den einzelnen Schularten in den Ländern der Bundesrepublik Deutschland zusammenhängen, kann in DESI aufgrund der ausschließlich national repräsentativen Stichprobe nicht beantwortet werden. Jenseits der fachspezifischen Differenzen in der Leistungsstreuung innerhalb der einzelnen Schularten, weisen die jeweiligen Varianzanteile in den Schülerkompetenzen, die auf Unterschiede zwischen Schulen derselben Schulart zurückgehen (14% bis 38% im Deutschen, 16% bis 35% im Englischen), auf die Potentiale hin, die für die pädagogische Qualitätsentwicklung und Förderung der Sprachkompetenzen der Schülerinnen und Schüler auf Schulebene genutzt werden können.

34.3 Institutionelle und professionelle Bedingungen des Sprachunterrichts aus Sicht der Fachkollegien Deutsch und Englisch

Mit der Ermittlung der individuellen und institutionellen Unterschiede in den sprachlichen Kompetenzen der Schülerinnen und Schüler auf der Basis der Varianzkomponenten konnte bisher Auskunft darüber gegeben werden, in welchem Ausmaß die Schülerkompetenzen zwischen den Schularten und den Einzelschulen variieren. Damit ist noch nichts über die *Qualität* der institutionellen Bedingungen für den Unterricht gesagt. Zur Erfassung des professionellen Handlungskontextes für den Sprachunterricht in Schulen wurden deshalb die Fachkollegien Deutsch und Englisch als professionelle Bezugsgruppe als Zielpopulation der Wahl herangezogen. Durch eine vergleichbare Ausbildung und berufliche Praxis verfügen die Mitglieder der Fachkollegien über einen gemeinsamen fachlichen Hintergrund. Sie verantworten darüber hinaus die Abstimmung der Ziele, Methoden und angestrebten Ergebnisse des Fachunterrichts. Weil DESI erstmals vollständige Fachkollegien für die Fächer Deutsch und Englisch in die Untersuchung einbezogen hat, gibt die Studie einen differenzierten Einblick in die professionellen Leitvorstellungen und kooperativen Praktiken der jeweils zentralen Referenzgruppe, auf die sich Lehrkräfte bei der Gestaltung des Sprachunterrichts beziehen.

Theoretischer Hintergrund für die Erfassung der institutionellen und professionellen Bedingungen des Sprachunterrichts sind die Konzepte und Konstrukte der Schulqualität und der professionellen Lerngemeinschaften aus der Schuleffektivitäts- und Schulentwicklungsforschung (Brophy/Good 1986; Ditton 2000; Fend 1998; Louis/Leithwood 1998; Reynolds u.a. 1996; Sammons/Thomas/Mortimore 1997a, 1997b; Scheerens/Bosker 1997; Slavin 1996; Stringfield 1994; Teddlie/Reynolds 2001; Wang/Haertel/Walberg 1993). Zunächst werden dort in einem formalen Modell für die Analyse schulischer Lernergebnisse die jeweils bedeutsamen Input-, Prozess- und Outputmerkmale auf den einzelnen Untersuchungsebenen Individuum, Klasse und Schule unterschieden, die auf der jeweils relevanten Ebene einen Erklärungsbeitrag für Unterschiede in den Lernergebnissen von Schülerinnen und Schülern leisten. Die Schulebene und Merkmale der Einzelschule erweisen sich nach den Ergebnissen verschiedener Metaanalysen im Vergleich zu individuellen Merkmalen der Schülerinnen und Schüler erwartungsgemäß als nicht so erklärungsmächtig für Schülerleistungen. Gleichwohl fanden sich vor allem in Studien aus dem anglo-amerikanischen Raum bedeutsame Unterschiede in den Schülerleistungen zwischen Einzelschulen, die auch unter Berücksichtigung sozialer Ausgangslagen der Schülerinnen und Schüler mit Schulmerkmalen variieren.

In einer Metaanalyse zahlreicher empirischer Primär- und Sekundäruntersuchungen zum schulischen Lehren und Lernen haben Wang/Haertel/Walberg (1993) das Konstrukt der Schulkultur als praktisch bedeutsamsten Faktor der Schuleffektivität auf Schulebene ermittelt. Zur Schulkultur gehören das Schulethos, das die gemeinsamen Normen und Werte bezüglich des Lehrens und Lernens in einer Schule beinhal-

tet, die Leistungsorientierung und die kooperativen Zielstrukturen und Praktiken der Lehrkräfte für den Unterricht. Im MACRO-Modell stellt Stringfield (1994) folgende Schulfaktoren eines effektiven Unterrichts als bedeutsam heraus: bedeutungsvolle, allgemein anerkannte Ziele (meaningful goals), Aufmerksamkeit für einen wirksamen Unterricht (attention to academic functioning), Koordination schulischer Ziele und Maßnahmen (coordination among programs), Rekrutierung und Fortbildung von Lehrkräften (recruitment of teachers, development of staff), lernförderliche Schulorganisation (organization of schools). In ihrem mehrbenenanalytischen Modell der Schuleffektivität untersuchen Scheerens/Bosker (1997) auf der Basis von Meta- und Sekundäranalysen als bedeutsame Schulfaktoren: die pädagogische Führung; Konsens, Kohäsion und Kooperation; Leistungsorientierung; Curriculumqualität und Lerngelegenheiten; Evaluation; Schulklima; Elterneinbeziehung. Dabei werden einige Konstrukte auch fachspezifisch operationalisiert: Leistungsorientierung als fachspezifische Leistungserwartung; Curriculumqualität als Abstimmung von Fachinhalten, Lehrbüchern und Unterrichtsmethoden; Konsens und Kooperation der Lehrkräfte hinsichtlich der pädagogischen und fachlichen Ziele sowie der Planung und Durchführung des Unterrichts; Evaluation als Überprüfung der Schülerleistungen und als Maßnahme der Selbstevaluation der Schule.

Mit dem Konzept der Professionellen Lerngemeinschaften greifen Louis/Leithwood (1998) im Rückgriff auf Kruse/Louis/Bryk (1995) Konstrukte des organisationellen und professionellen Lernens auf. Einen Beitrag zur Steigerung schulischer Effektivität leisten danach die Faktoren: gemeinsame Normen und Wertvorstellungen (shared norms and values), die wechselseitige Verpflichtung auf das Lernen der Schülerinnen und Schüler (focus on student learning), eine regelmäßige Diskussion und Kooperation der Lehrkräfte zu Fragen des Curriculums, des Unterrichts und der Leistungsbeurteilung von Schülerinnen und Schülern (reflective dialog, collaboration) sowie eine offene Diskussion über Unterrichtspraktiken unter den Lehrkräften (de-privatization of practice). Gemeinsame Ziele und Werte des Lehrerkollegiums und eine kooperative Umsetzung in die Unterrichtspraxis haben auch Mortimore u.a. (1988) und Mortimore (1994) mit Bezug auf ihre längsschnittliche Grundschulstudie als Voraussetzung für konsistentes Lehrerhandeln und Bedingung für Schuleffektivität hervor.

Hoy/Tarter/Hoy (2006) identifizieren in einer Stichprobe von 96 Sekundarschulen eines Bundesstaates aus dem mittleren Westen der USA den akademischen Optimismus im Lehrkörper (academic optimism) als organisationelles Merkmal der Schule, das auf Schulebene auch unter Kontrolle des sozialen Hintergrunds und des Vorwissens für Schulunterschiede in den Schülerleistungen sowohl in Mathematik und Naturwissenschaften als auch im Bereich der Sprache bedeutsam ist. Der akademische Optimismus setzt sich als latentes Konstrukt zusammen aus der Betonung akademischer Ansprüche und Praktiken (academic emphasis), kollektiver Selbstwirksamkeit (collective efficacy) und dem Vertrauen der Lehrerkollegien in das Engagement von Schülerinnen und Schülern sowie Eltern (faculty trust in students and parents). Der akademische Optimismus zeigt dort einen positiven Effekt

auf die Schülerleistungen, und zwar direkt und indirekt über den sozioökonomischen Status und die Ausgangsleistungen der Schülerinnen und Schüler.

Auf der Basis einer Längsschnittstudie in flämischen Sekundarschulen (Belgien) wurde die Entwicklung von fachlichen und überfachlichen Lernergebnissen von Schülerinnen und Schülern hinsichtlich mehrerer Zielkriterien (Schülerleistung in Mathematik, Niederländisch, affektive und soziale Lernergebnisse) und im Zusammenhang mit verschiedenen schulischen und unterrichtlichen Bedingungsfaktoren untersucht. Zunächst wurde auch dort das Ausmaß der institutionellen und individuellen Leistungsunterschiede ermittelt. Opdenakker und van Damme (2000) berichten am Ende des ersten Sekundarschuljahres relativ große Unterschiede zwischen Klassen (33%) und zwischen Schulen (23%) in den Sprachkompetenzen der Schülerinnen und Schüler im Niederländischen. Durch die Einführung typischer Erklärungsmerkmale aus der Schuleffektivitätsforschung auf Schülerebene (Intelligenz) und auf Schulebene (u. a. Lehrerkooperation, Disziplin- und Leistungsorientierung, geordnete Lernumgebung/Klassenführung, Differenzierung im Unterricht) in das Erklärungsmodell reduzieren sich die Unterschiede in den Schülerleistungen zwischen Schulen und Klassen erheblich. Gleichwohl zeigen einzelne Schulmerkmale wie die geordnete Lernumgebung und die Lehrerkooperation einen positiven Effekt auf die Sprachleistungen am Ende der ersten Sekundarschulklasse, die Differenzierung im Unterricht einen negativen Effekt. In einer weiteren Mehrebenenanalyse an demselben Datensatz, in der insbesondere Kompositionseffekte auf Klassenebene auf die Schülerleistungen überprüft wurden, zeigen sich die Zusammenhänge zwischen Lehrerkooperation und sprachlichen Leistungen nicht mehr. De Fraine et al. (2003) können die Leistungsunterschiede im Niederländischen nicht mehr durch schulische Prozessmerkmale bzw. Erziehungspraktiken der Lehrkräfte (i. e. Betonung individueller Förderung, Förderung leistungsstarker bzw. leistungsschwacher Schüler, Beratungen unter Lehrkräften über Methoden und Schüler, Rückmeldungen an die Schülerinnen und Schüler, geordnete Lernumgebung) erklären, wenn sie mehr Individualmerkmale als die Intelligenz der Schülerinnen und Schüler (i. e. Geschlecht, sozioökonomischer Status, Leistungsmotivation, sprachliche Ausgangsleistungen) in das Modell einführen und zusätzlich das kognitive Ausgangsniveau und den Mädchenanteil auf Klassenebene kontrollieren. Im Ergebnis zeigen sich deutliche Kompositionseffekte des kognitiven Leistungsniveaus und des Mädchenanteils in den Klassen auf die Leistungen im Niederländischen. Überdies stellen de Fraine et al. (2003) differentielle Effekte auf Klassenebene fest. In der Gruppe der Schülerinnen und Schüler mit niedrigen Ausgangsleistungen gibt es Klassen, in denen sich die sprachlichen Leistungen stärker verbessern als in anderen. Diese Unterschiede können die Autoren jedoch nicht durch ihre Schul- und Unterrichtsfaktoren erklären. Die Bedeutung schulischer Normen und Praktiken für die Schülerleistungen variieren in den berichteten Analysen mit dem jeweils zugrunde gelegten Erklärungsmodell. Kompositionelle Effekte in der kognitiven Zusammensetzung der Schülerschaft können die Effekte von Schul- und Unterrichtsmerkmalen überlagern. Um diese Effekte auseinander halten zu können, sind in extern differenzierenden Schulsystemen

wie Deutschland die Schularten als institutionelle Bedingungsfaktoren und die Schülerzusammensetzung bei der Prüfung der Bedeutsamkeit schulischer Gestaltungsmerkmale für Schülerleistungen zu berücksichtigen.

Im deutschsprachigen Raum hat Fend in seinen großen Leistungsstudien in den Jahren 1978/79 – neben allgemeinen Lehrereinstellungen und einschlägigen Schülermerkmalen – Einschätzungen der Lehrkräfte zu Fragen der Leistungsorientierung, Kooperation und Schülerorientierung als Prädiktoren der Schulqualität für Schülerleistungen erfasst (vgl. Fend 1998). Auf diese Konstrukte der Schulqualität greift auch Ditton (2000) in seinem Modell der Schul- und Unterrichtsqualität mit Bezug auf das QAIT-MACRO-Modell von Slavin (1996) und Stringfield (1994) zurück: Schulkultur als gemeinsame Zielvorstellungen bezüglich der Aufgaben und Verantwortlichkeiten; Kooperation und Koordination für einen geregelten Schul- und Unterrichtsbetrieb und einen professionellen Erfahrungsaustausch; Schulmanagement als gemeinsames Aufgabenverständnis und geregelte Aufgabenverteilung; Personalentwicklung durch interne Kooperation und Fort- und Weiterbildung. Innerhalb der aktuellen internationalen Schulleistungsstudien werden Schulmerkmale vor allem zu deskriptiven Zwecken erhoben. Lehrerk Kooperation wird beispielsweise über Fragen nach der Häufigkeit von Lehrerkonferenzen, Gesprächen über pädagogische Themen oder nach dem Zusammenhalt unter den Lehrkräften erfasst. So beurteilen in PISA 2000 die Schulleitungen das Arbeitsethos sowie Konsens und Kohäsion im Kollegium in den PISA-Schulen in Deutschland mehrheitlich positiv (Weiß/Steinert 2001). In PISA 2003 zeigen sich bei Konsens und Kooperation unter Mathematiklehrkräften deutliche Schulartunterschiede (Senkbeil u.a. 2004). In IGLU wurde nach der offiziellen Absicherung der Lehrerk Kooperation gefragt: Sie ist in Deutschland wie auch in anderen Ländern selten gegeben (Radisch/Steinert 2005). Auch wenn diese Befunde wichtige Aspekte der internen Arbeitsorganisation in den Blick nehmen, sind die Anforderungen an eine kohärente und effektivitätsfördernde Lernumgebung in Schulen inhaltlich und methodisch unscharf konzeptualisiert, dass sich keine Zusammenhänge mit den untersuchten Zielkriterien zeigen (vgl. Steinert u.a. 2006; Creemers/Kyriakides 2006).

Schulmerkmale können für Unterricht und Lernen bedeutsam sein, wenn sie zu einem konsistenten Lehrerhandeln und einer kohärenten Lernumgebung beitragen. Um Anhaltspunkte dafür zu gewinnen, ob und inwieweit die auf die Sprachkompetenzen und den Sprachunterricht bezogenen Leitvorstellungen und Praktiken der Fachkollegien für die Sprachkompetenzen von Bedeutung sind, wurden in DESI die sprachbezogenen Einschätzungen und zentrale Formen und Inhalte kooperativen Handelns in den Fachkollegien Deutsch und Englisch erhoben. Da die Leitvorstellungen in den Fachkollegien vermutlich die Zielstruktur des Deutsch- bzw. Englischunterrichts beeinflussen, wurden die Mitglieder der Fachkollegien nach ihren Leistungserwartungen und unterrichtlichen Schwerpunktsetzungen gefragt. In der Frage nach den Leistungen, die die Mehrheit der Schülerinnen und Schüler einer Schule am Ende der neunten Jahrgangsstufe voraussichtlich erbringen werden, sollten die Lehrkräfte Voraussetzungen und die Erträge kumulativen

Lernens in Bezug auf die in DESI getesteten Kompetenzen am Ende der neunten Jahrgangsstufe einschätzen und damit die Leistungsfähigkeit ihrer Schülerinnen und Schüler auf Schulebene antizipieren. Mit der Frage nach der Wichtigkeit einzelner sprachlicher Kompetenzen im Unterricht wurde die normative Zielstruktur bzw. das gemeinsam geteilte Ziel- und Aufgabenverständnis des Sprachunterrichts erfasst. Mit der Frage nach der Kooperation und dem Austausch im Fachkollegium wurden die Praktiken erhoben, mit denen das Fachkollegium die inhaltlichen, didaktischen und diagnostischen Fragen des Sprachunterrichts, das Unterrichtshandeln und die professionelle Entwicklung der Lehrkräfte untereinander abstimmt. Die einzelnen verwendeten Skalen gehen aus der Übersicht in Tabelle 34.2 hervor. Für jede Skala werden ein Beispielitem, der Mittelwert, die Standardabweichung und die Anzahl der Items pro Skala sowie die interne Konsistenz der Skala angegeben. Die Verteilungen und Skalenkennwerte wurden auf Individualebene, d.h. auf Basis der befragten Lehrkräfte der jeweiligen Fachkollegien, ermittelt.

In den Abbildungen 34.5 bis 34.8 sind die Mittelwerte der professionellen Leitvorstellungen und kooperativen Praktiken der Fachkollegien Deutsch und Englisch nach Schularten dargestellt. Danach korrespondieren die Unterschiede in den Leistungseinschätzungen und in der Fokussierung auf einzelne sprachliche Kompetenzen in den Fachkollegien mit den schulartspezifischen Unterschieden in den Sprachkompetenzen der Schülerinnen und Schüler (vgl. Abbildung 34.1 und 34.2). Überdurchschnittliche Leistungen erwarten die Fachkollegien Deutsch und Englisch in den Gymnasien und in den Realschulen, unterdurchschnittliche in den übrigen Schularten (Abbildung 34.5). Die Fachkollegien an Integrierten Gesamtschulen erwarten bei den Englischkompetenzen in Relation zu den anderen Schularten niedrigere Leistungen als es sich im tatsächlichen Leistungsniveau abzeichnet, während die Fachkollegien aus den Schulen mit mehreren Bildungsgängen im Vergleich zu den anderen Schularten nicht so niedrige Leistungen im Englischen annehmen, wie ihre Schülerinnen und Schüler im Durchschnitt erzielen. Analog zur schulartspezifischen Verteilung der Schülerkompetenzen zeigen sich bei den Leistungserwartungen der Fachkollegien hinsichtlich des Englischen größere Schulartunterschiede als hinsichtlich des Deutschen. Die Schulart erklärt 16% der Variation in den Leistungserwartungen im Fachkollegium Deutsch und 44% der Variation in den Leistungserwartungen im Fachkollegium Englisch².

2 Die Effektgröße η^2 gibt den Varianzanteil der abhängigen Variablen (hier: Einschätzungen des Fachkollegiums) an, der durch die unabhängige Variable (hier: Schulart) erklärt wird.

Tabelle 34.2: Leitvorstellungen und kooperative Praktiken im Fachkollegium Deutsch und Englisch (Skalen, Beispielitems, Mittelwert (MW), Standardabweichung (SD), Anzahl der Items und interne Konsistenz (α)).

Skalen und Beispielitems	MW	SD	Items	α
Alle Items haben ein vierstufiges Antwortformat von 1-4. Die Skalenwerte wurden als Mittelwerte gebildet.				
Professionelle Leitvorstellungen im Fachkollegium				
Leistungserwartung im Fach Deutsch	2.54	0.45	7	.76
Leistungserwartung im Fach Englisch	2.82	0.53	9	.91
<i>Wenn ich in unserer Schule eine neue Klasse/Kurs übernehme, kann ich im Fach Deutsch/Englisch auf soliden Vorkenntnissen der Schülerinnen und Schüler aufbauen.</i>				
Fokussierung auf sprachliche Kompetenzen				
Schriftliche Kommunikationsfähigkeit (Deutsch)	3.55	0.36	5	.67
Schriftliche Kommunikationsfähigkeit (Englisch)	3.43	0.38	5	.71
Kooperative Praktiken im Fachkollegium				
Fachinhaltliche Kooperation				
Fachinhaltliche Kooperation im Fachkollegium Deutsch	2.66	0.58	9	.87
Fachinhaltliche Kooperation im Fachkollegium Englisch	2.70	0.53	9	.85
<i>Im Fachkollegium tausche ich mit Kolleg/innen Aufgabenstellungen für Klassen-, Kursarbeiten, Schulaufgaben aus.</i>				
Fachdidaktische Kooperation				
Fachdidaktische Kooperation im Fachkollegium Deutsch	2.36	0.55	6	.76
Fachdidaktische Kooperation im Fachkollegium Englisch	2.33	0.55	6	.76
<i>Im Fachkollegium beraten wir die Möglichkeiten individueller Förderung von Schülerinnen und Schülern.</i>				
Diagnostisch-methodische Kooperation				
Diagnostisch-methodische Kooperation im Fachkollegium Deutsch	2.09	0.53	6	.75
Diagnostisch-methodische Kooperation im Fachkollegium Englisch	2.13	0.50	7	.71
<i>Im Fachkollegium beraten wir, wie wir spezifische Stärken und Schwächen der Schülerinnen und Schüler besser erkennen.</i>				
Zielkoordination				
Zielkoordination im Fachkollegium Deutsch	2.92	0.62	5	.78
Zielkoordination im Fachkollegium Englisch	2.77	0.59	5	.78
<i>Wir gestalten den Deutschunterricht / Englischunterricht nach gemeinsamen pädagogischen Zielen.</i>				
Austausch von Materialien				
Austausch von Materialien im Fachkollegium Deutsch	2.92	0.60	3	.73
Austausch von Materialien im Fachkollegium Englisch	2.91	0.59	3	.72
<i>Mit Fachkolleg/innen tausche ich regelmäßig Unterrichtsmaterial aus.</i>				
Austausch über Unterrichtserfahrungen				
Austausch über Unterrichtserfahrungen im Fachkollegium Deutsch	2.62	0.64	3	.71
Austausch über Unterrichtserfahrungen im Fachkollegium Englisch	2.67	0.61	3	.71
<i>Die Kommunikation im Fachkollegium trägt zum offenen Erfahrungsaustausch über die eigene Unterrichtspraxis bei.</i>				

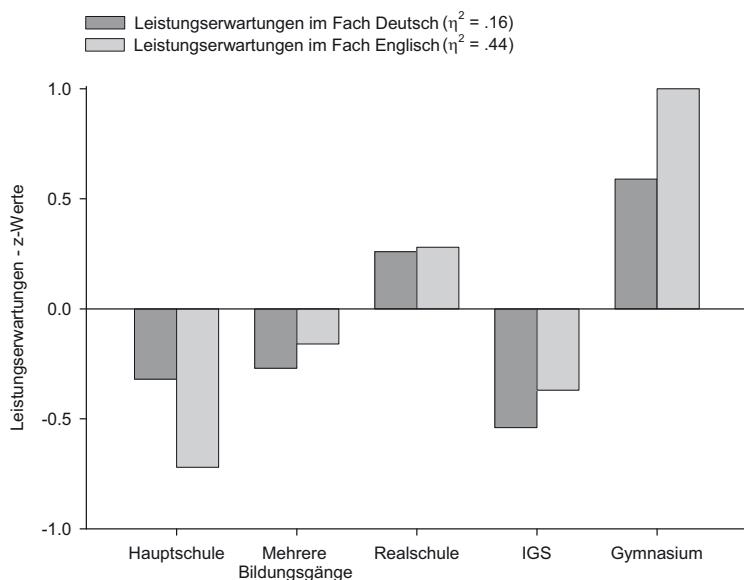


Abbildung 34.5: Einschätzung der Schülerkompetenzen am Ende der neunten Jahrgangsstufe durch die Fachkollegien Deutsch und Englisch nach Schularten (z-Werte).

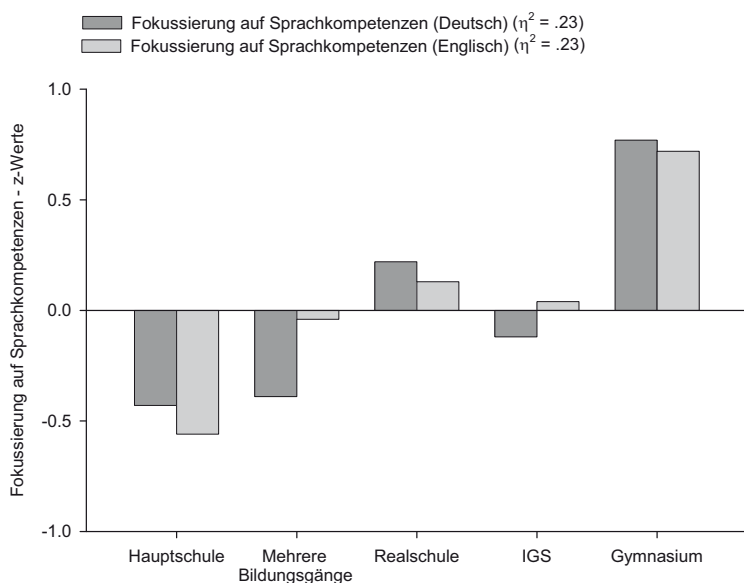


Abbildung 34.6: Fokussierung auf Sprachkompetenzen im Unterricht durch die Fachkollegien Deutsch und Englisch nach Schularten (z-Werte).

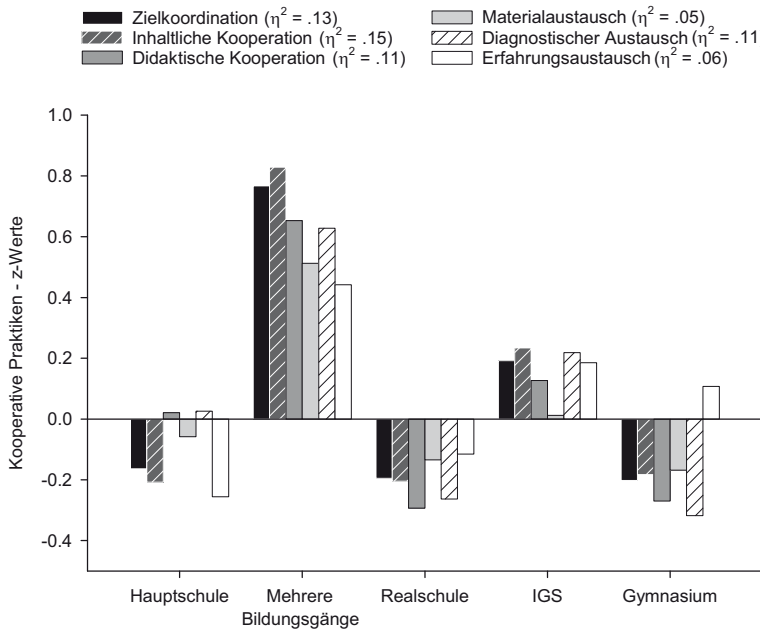


Abbildung 34.7: Kooperative Praktiken im Fachkollegium Deutsch nach Schularten (z-Werte).

In den Gymnasien legen beide Fachkollegien auf die einzelnen sprachlichen Kompetenzen im Deutsch- und im Englischunterricht besonderes Gewicht; in den Hauptschulen ist dies weit weniger der Fall. In der Fokussierung auf die Sprachkompetenzen liegen der anderen Schularten näher beieinander. Insgesamt erklärt die Schulart jeweils 23% der Varianz in den normativen Leitvorstellungen der Fachkollegien Deutsch und Englisch (vgl. Abbildung 34.6).

In den Fachkollegien Deutsch und Englisch zeichnen sich im Hinblick auf die einzelnen Formen und Inhalte der kooperativen Praktiken ähnliche Präferenzen ab (vgl. Tabelle 34.2): Am häufigsten kommt in den Fachkollegien der Austausch von Lehr- und Unterrichtsmaterialien und die Abstimmung von Zielsetzungen vor. Die Kooperation zu Fachinhalten und der Austausch von Unterrichtserfahrungen folgen an zweiter Stelle. Weniger praktiziert wird die Kooperation zu didaktischen Fragen und zu methodisch-diagnostischen Fragen. Selten kommt es vor, dass die Lehrkräfte wechselseitig im Unterricht hospitieren. Dieses Häufigkeitsmuster kooperativen Handelns deckt sich mit den Befunden aus anderen Studien zur Lehrerverkooperation, die darlegen, dass die einzelnen Formen und Inhalte der Lehrerverkooperation hinsichtlich der Zielinterdependenz und Wechselseitigkeit der Bezugnahme unterschiedlich voraussetzungsvoll sind und insofern in ihrer Nutzung zwischen einzelnen Schulen variieren (vgl. Rolff 1980, 1991; Gräsel/Fußangel/Pröbstel 2006; Steinert u.a. 2006). Wie in Abbildung 34.7 und 34.8 zu sehen, sind die Schulartunterschiede im Kooperationsverhalten zwischen den Fachkollegien in Deutsch und Englisch ähnlich.

Die schulartspezifischen Unterschiede sind für die kooperativen Praktiken auf der einen und den Leistungseinschätzungen und der Fokussierung auf Sprachkompetenzen auf der anderen Seite deutlich verschieden (vgl. Abbildung 34.5 und 34.7). Die Kooperation in den Schulen mit mehreren Bildungsgängen und in den Integrierten Gesamtschulen ist überdurchschnittlich ausgeprägt und dürfte Folge der Fachleistungsdifferenzierung und des erhöhten Koordinierungsbedarfs in Schulen sein, die mehrere Bildungsgänge anbieten. In den Schulen des traditionell gegliederten Schulwesens sind die einzelnen Kooperationspraktiken durchschnittlich bis unterdurchschnittlich ausgeprägt.

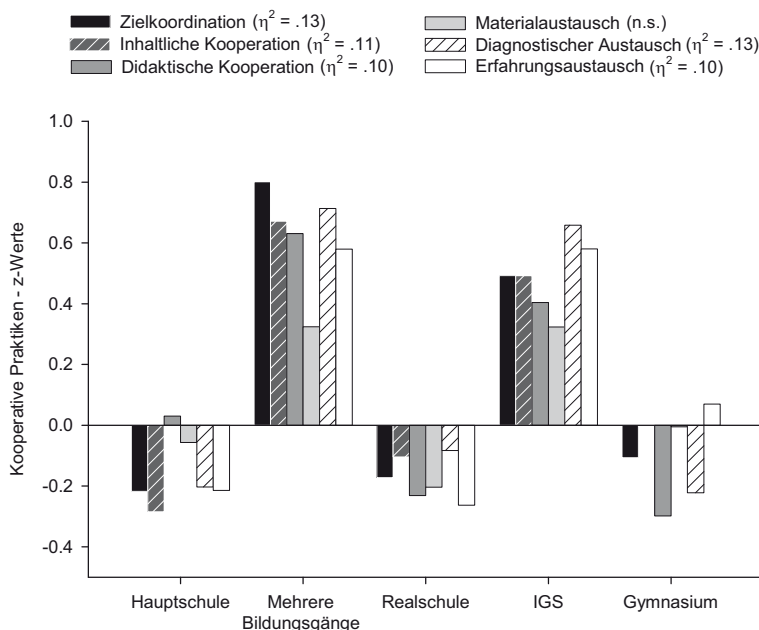


Abbildung 34.8: Kooperative Praktiken im Fachkollegium Englisch nach Schularten (z-Werte).

Die Schulartunterschiede im kooperativen Handeln sind im Vergleich zu den Leitvorstellungen der Fachkollegien zwar kleiner, aber in den meisten Aspekten durchaus bedeutsam. Besonders ausgeprägt sind die Unterschiede zwischen den Schularten in den kooperativen Praktiken der Fachkollegien Deutsch und Englisch bei der Zielkoordination, der fachinhaltlichen Kooperation und dem Austausch über die Diagnose und Förderung von Schülerleistungen. Jenseits der schulartspezifischen Ausprägungen gibt es in den professionellen Leitvorstellungen und kooperativen Praktiken der Fachkollegien als institutionelle Bedingungen des Sprachunterrichts eine deutliche Variation zwischen den Schulen, die im Folgenden auf ihren Zusammenhang mit den sprachlichen Kompetenzen der Schülerinnen und Schüler untersucht wird.

34.4 Zusammenhänge der professionellen Leitvorstellungen und kooperativen Praktiken der Fachkollegien mit den Sprachkompetenzen der Schülerinnen und Schüler

Die in Abschnitt 34.2 dargestellten institutionellen Varianzkomponenten zeigen das Ausmaß an, in welchem die sprachlichen Leistungen der Schülerinnen und Schüler auf der Ebene der Einzelschule einzelschulischen Gestaltungsparametern zugänglich sein können. In diesem Abschnitt wird deshalb geprüft, ob die von der Schuleffektivitäts- und Schulentwicklungsforschung identifizierten Prozessmerkmale auf Schulebene für die Sprachkompetenzen der Neuntklässler von Bedeutung sind. Die in DESI ermittelten Zusammenhänge werden dann den empirischen Befunden anderer Schulleistungsstudien gegenübergestellt, um Anhaltspunkte für die Generalisierbarkeit der Befunde z.B. über Fächer hinweg zu erhalten. Denn die leistungsfördernde Bedeutung der Schulmerkmale ist keineswegs so eindeutig, wie dies häufig angenommen wird. Aufgrund der Unterschiede in der Konzeptualisierung der verschiedenen Studien variieren Stärke und Richtung der gefundenen Zusammenhänge. Die Studiendesigns unterscheiden sich im Hinblick auf die untersuchten Länder, Schulstufen und Fächer. Auch Differenzen in der Stichprobenziehung und im Design haben Folgen für die Generalisierbarkeit der einzelnen Befunde. In querschnittlichen Untersuchungen können ohne Kontrolle der Ausgangsleistungen die Effekte von Schulen und von Schulmerkmalen auf Schülerleistungen überschätzt werden. Bei Nichtberücksichtigung der Mehrebenenstruktur von Schulleistungsdaten werden die tatsächlichen Zusammenhänge zwischen den einzelnen Aggregateinheiten und den individuellen Schülerleistungen nicht korrekt geschätzt. Insofern ist es keine Überraschung, wenn Scheerens/Bosker (1997) in ihren Metaanalysen feststellen, dass die Zusammenhänge zwischen ausgewählten Schuleffektivitätsmerkmalen und Schülerleistungen in den verschiedenen Studien nicht konsistent sind und die Schwelle statistischer Signifikanz häufig verfehlen. Für die Lehrerkooperation und die Leistungsorientierung als Schulmerkmale finden sie beispielsweise sowohl positive als auch negative oder gar keine Zusammenhänge mit der Schülerleistung. Eindeutig positive Effekte zeigen beide Merkmale nur in den USA. Auch Fend (1998) berichtet aus seinen großen Leistungsuntersuchungen keine durchgängigen und gleichgerichteten Zusammenhänge zwischen Leistungsorientierung, Kooperation und den Schülerleistungen. Je nach Schulart finden sich positive oder keine Zusammenhänge zwischen Leistungsorientierung und Schülerleistung sowie positive, negative oder keine Zusammenhänge zwischen Lehrerkooperation und Schülerleistung.

In DESI konzentriert sich die Analyse der institutionellen Bedingungen für den Fachunterricht auf die professionellen Leitvorstellungen und kooperativen Praktiken in den Fachkollegien Deutsch und Englisch. Diese wurden einmalig am Ende der neunten Jahrgangsstufe erhoben, ausgewählte sprachliche Kompetenzen hingegen zu Beginn und am Ende der neunten Jahrgangsstufe. Die ermittelten Zusammenhänge zwischen den professionellen Leitvorstellungen und kooperativen Praktiken der Fach-

kollegien als Parameter der Schulqualität und den Schülerkompetenzen am Ende der neunten Jahrgangsstufe erlauben Aussagen darüber, ob Schulen – unter Berücksichtigung zentraler individueller Merkmale der Schülerinnen und Schüler und unter Berücksichtigung der Zusammensetzung der Schülerschaft auf Schulebene – „erwartungswidrig“ hohe oder niedrige sprachliche Kompetenzen und Kompetenzzuwächse aufweisen. Ein „erwartungswidriges“ Kompetenzniveau erzielen Schulen, wenn das Leistungsergebnis einer Schule sich signifikant von dem Ergebnis unterscheidet, das aufgrund der individuellen Voraussetzungen ihrer Schülerinnen und Schüler und der Zusammensetzung ihrer Schülerschaft zu erwarten gewesen wäre.

Für die Analyse dieser Zusammenhänge wurden auf Individualebene das Geschlecht (Referenzkategorie: Mädchen), die kognitiven Grundfähigkeiten, der sozioökonomische Status, die Erstsprache (Referenzkategorie: deutsch) und die Teilnahme am bilingualen Unterricht als kognitive und soziale Hintergrundmerkmale der Schülerinnen und Schüler kontrolliert, auf Schulebene die Schulart, der Mädchenanteil und die kognitive, soziale und sprachliche Zusammensetzung der Schülerschaft sowie die Zugehörigkeit zu einer bilingualen Schule (vgl. Tabelle 34.3). Weil die Hauptschulen, Schulen mit mehreren Bildungsgängen und integrierten Gesamtschulen Schülerinnen und Schüler mit ähnlichen Ausgangslagen beschulen und einzeln eine geringe Fallzahl aufweisen, wurden diese Schularten zu einer Gruppe zusammengefasst. Referenzkategorie waren die Realschulen, die den übrigen Schularten und den Gymnasien gegenübergestellt wurden. Angesichts des Oversampling der bilingualen Schulen in DESI und der erwarteten Wirksamkeit des bilingualen Sachfachunterrichts wurden auch das schulische Angebot und die individuelle Teilnahme am bilingualen Unterricht zusätzlich kontrolliert.

Zusammenfassung der Indikatoren zu übergeordneten Faktoren

Als Indikatoren für eine zielorientierte und fachlich abgestimmte Lernumgebung wurden die auf Schulebene aggregierten Urteile der Fachlehrerinnen und -lehrer zu ihren professionellen Leitvorstellungen und kooperativen Praktiken verwendet. Diese aggregierten Werte wurden als Prädiktoren der Sprachkompetenzen auf Schulebene in ein Zweiebenenstrukturgleichungsmodell eingeführt. Die entsprechenden Analysen wurden mit dem Programm Mplus berechnet (Muthén/Muthén 1998-2004; vgl. Kap. 3). Die in Tabelle 34.2 aufgeführten Skalen, die sich auf ähnliche Bereiche der Zielvorstellungen und Praktiken der Fachkollegien beziehen, sind untereinander korreliert. Ziel der Analysen war es daher zunächst, eine Struktur zu finden, in der die Einschätzungen der Fachkollegien zu einer geringeren Anzahl übergeordneter Dimensionen zusammengefasst werden können. Die Zusammenhänge der Indikatoren auf Schulebene lassen sich insgesamt durch ein dreifaktorielles Strukturmodell gut abbilden. In diesem Modell, das sowohl für die Fachkollegien Deutsch als auch Englisch gute Modellanpassungen zeigt, werden die Dimensionen *Leistungseinschätzungen der Sprachkompetenzen* und *Fokussierung*

auf *Sprachkompetenzen* jeweils durch eine einzelne Skala repräsentiert. Die verschiedenen Formen und Inhalte der Kooperation im Fachkollegium konnten hingegen zu einem übergeordneten Faktor zusammengefasst werden. Diese Dimension *Kooperative Praktiken im Fachkollegium* wird durch sechs manifeste Indikatoren definiert: Zielkoordination, Kooperation zu fachinhaltlichen, fachdidaktischen und diagnostischen Fragen, Austausch von Materialien und von Erfahrungen.

Die Dimensionalität der Einschätzungen der Fachkollegien entspricht zum einen den globalen Merkmalen aus der Schulqualitätsforschung „Leistungsorientierung“ bzw. „Fokus auf das Lernen der Schüler“ und „Kooperation“, zeigt aber auch als Konfiguration Zusammenhänge auf, die nicht unbedingt zu erwarten gewesen waren. Die Antizipation und die normativen Vorstellungen der Fachkollegien bezüglich der sprachlichen Kompetenzen der Schülerschaft ihrer Schule stellen sowohl im Deutschen als auch im Englischen zwei eigenständige Dimensionen dar, die nur moderat miteinander korrelieren ($r = .32$ im Fachkollegium Deutsch und $r = .53$ im Fachkollegium Englisch). Fachkollegien, die die sprachliche Leistungsfähigkeit ihrer Schülerschaft höher einschätzen, haben tendenziell auch ein höheres Anspruchsniveau bezüglich der einzelnen Sprachkompetenzen. Der Zusammenhang ist allerdings nicht so eng, dass nicht auch Fachkollegien, die die Leistungsstärke ihrer Schülerinnen und Schüler weniger hoch einschätzen, im Unterricht auf die einzelnen sprachlichen Kompetenzen besonderen Wert legen würden.

Die verschiedenen Formen und Inhalte der Kooperation unter den Fachlehrkräften lassen sich entgegen den Differenzierungen der Schulentwicklungsforschung auf ein gemeinsames latentes Konstrukt „Kooperation im Fachkollegium“ zurückführen. Dieses schließt über den Austausch und die Zusammenarbeit zu fachlichen, didaktischen und diagnostischen Fragen des Unterrichts hinaus auch Fragen der Koordination von Zielen bzw. Anforderungen und Ansätze zur gegenseitigen Beurteilung des Unterrichtshandelns ein. Zwischen dem Faktor Kooperation auf der einen und den Leistungseinschätzungen und der Fokussierung auf Sprachkompetenzen als normative Leitvorstellungen der Fachkollegien Deutsch und Englisch auf der anderen Seite besteht kein signifikanter Zusammenhang. Dies ist vor allem deshalb hervorzuheben, weil ein gemeinsames Ziel- und Aufgabenverständnis nicht nur Voraussetzung für Kooperation, sondern auch für deren Wirksamkeit ist (vgl. Gräsel/Fußangel/Pröbstel 2006). Die Unabhängigkeit der kooperativen Praktiken von den Vorstellungen der Fachkollegien über die erwarteten und geforderten Sprachkompetenzen bei den Schülerinnen und Schülern deutet darauf hin, dass die Fachlehrkräfte aufeinander Bezug nehmen, um beispielsweise die Unterrichtsvorbereitung effizient zu gestalten, diese Aktivitäten jedoch wenig mit den Zielen und Anforderungen an den kumulativen Kompetenzerwerb der Schülerinnen und Schüler koppeln.

Vorhersage der sprachlichen Kompetenzen

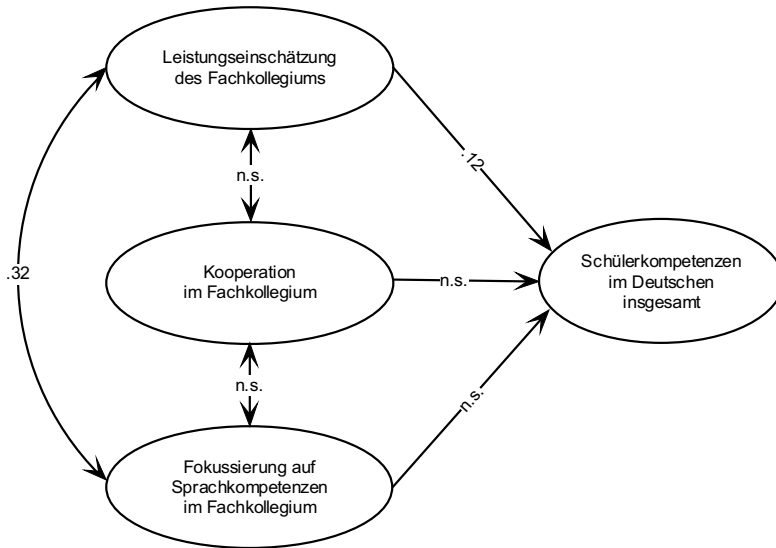
Die Konstrukte *Leistungseinschätzungen*, *Kooperation* und *Fokussierung auf Sprachkompetenzen innerhalb der Fachkollegien* wurden in weiteren Analysen als Prädik-

toren für die sprachlichen Kompetenzen der Schülerinnen und Schüler auf Schulebene verwendet. In diesen Modellen wurde untersucht, welchen Erklärungsbeitrag die aus den Berichten der Fachkollegien gebildeten Konstrukte der professionellen Leitvorstellungen und kooperativen Praktiken für die Sprachkompetenzen der Schülerinnen und Schüler leisten, wenn die in Tabelle 34.3 aufgeführten Variablen für die individuellen Schülervoraussetzungen und die Zusammensetzung der Schülerschaft auf Schulebene sowie die Schulart kontrolliert werden. Als abhängige Variablen wurden die Plausible Values für die Gesamtleistungen in Deutsch und Englisch sowie für ausgewählte sprachliche Kompetenzen die Differenzen zwischen den Plausible Values der Sprachleistungen zu Beginn und am Ende der neunten Jahrgangsstufe verwendet (vgl. Kap. 3).

Tabelle 34.3: Modellierung des Zusammenhangs zwischen den professionellen Leitvorstellungen und kooperativen Praktiken in den Fachkollegien Deutsch und Englisch und Sprachkompetenzen der Schülerinnen und Schüler (Kontrollvariablen sind normal, die interessierenden Prädiktoren kursiv gedruckt).

Kontrollvariablen und Prädiktoren	Kriterien
Individualebene <ul style="list-style-type: none"> - Geschlecht - Kognitive Grundfähigkeiten - Sozioökonomischer Status: HISEI - Erstsprache - Teilnahme am bilingualen Unterricht 	<ul style="list-style-type: none"> - Gesamtleistung: Deutsch - Gesamtleistung: Englisch - Zuwachs Deutsch: Schreiben; Sprachbewusstheit - Zuwachs Englisch: Textrekonstruktion; Hörverstehen
Schulebene <ul style="list-style-type: none"> - Schularten - Mädchenanteil - Kognitive, soziale und sprachliche Zusammensetzung der Schülerschaft - Zugehörigkeit zu einer bilingualen Schule - <i>Leistungseinschätzungen des Fachkollegiums</i> - Kooperative Praktiken im Fachkollegium - Fokussierung auf Sprachkompetenzen im Fachkollegium 	

In den folgenden Abbildungen sind aus Gründen der Übersichtlichkeit als Ergebnisse der Zweiebenenstrukturgleichungsmodelle nur die um die individuellen und institutionellen Ausgangslagen bereinigten standardisierten Effekte der Konstrukte *Leistungseinschätzungen*, *Fokussierung auf Sprachkompetenzen* und *Kooperation im Fachkollegium* auf die Schülerkompetenzen im Deutschen und im Englischen auf Schulebene dargestellt. Die standardisierten Regressionskoeffizienten β zeigen den Netto-Effekt der Schulfaktoren auf die individuellen Schülerleistungen an.



Modellfit: $\chi^2 = 158.713$; $df = 80$; CFI = .97; RMSEA = .01; SRMR_{between} = .08; $R^2_{\text{between}} = .86$.

Abbildung 34.9: Zusammenhang zwischen den professionellen Leitvorstellungen und kooperativen Praktiken im Fachkollegium Deutsch und den Schülerkompetenzen im Deutschen unter Berücksichtigung der in Tabelle 34.3 aufgeführten Kontrollvariablen.

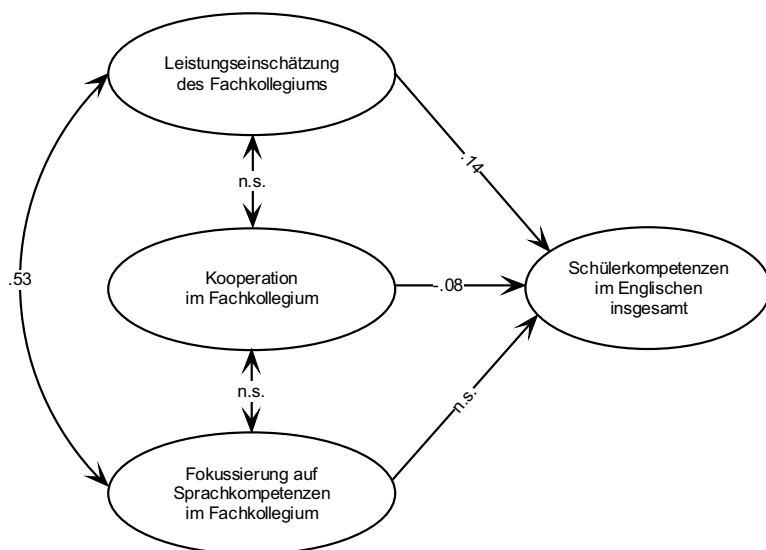
Aus Abbildung 34.9 geht hervor, dass Schülerinnen und Schüler an Schulen, an denen das Fachkollegium Deutsch – selbstverständlich ohne Kenntnis der Testergebnisse – die Erwartung teilt, dass ihre Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe die in DESI gestellten Anforderungen an die Sprachkompetenzen im Deutschen mehrheitlich erfüllen, unter Berücksichtigung individueller Lernvoraussetzungen und unter Berücksichtigung der Schulart und der kognitiven, sprachlichen und sozialen Zusammensetzung der Schülerschaft höhere Kompetenzen zeigen ($\beta = .12$). Die gemeinsame Fokussierung des Deutschunterrichts z.B. auf (schrift)sprachliche Korrektheit und Angemessenheit zeigt jedoch keinen signifikanten Zusammenhang mit den Schülerkompetenzen am Ende der neunten Jahrgangsstufe. Die normativen Zielvorstellungen in den Fachkollegien sind offenbar allein nicht hinreichend, um für die kumulativen Lernergebnisse am Ende der neunten Jahrgangsstufe bedeutsam zu sein.

Die kooperativen Praktiken des Fachkollegiums Deutsch weisen ebenso wie die normativen Leitvorstellungen keinen signifikanten Effekt auf die Sprachkompetenzen der Schülerinnen und Schüler auf. Die einzelnen Formen und Inhalte der Kooperation mögen insgesamt zu geregelten Unterrichtsabläufen und zu einer effizienten Unterrichtsvor- und -nachbereitung beitragen, erhöhen aber nicht die Effektivität des Unterrichts im Hinblick auf die getesteten Sprachkompetenzen. In separaten Mehrebenenanalysen zeigen zwei Einzelskalen zur Kooperation in den Fachkollegien Deutsch sogar einen negativen Zusammenhang mit den Schülerkompetenzen am Ende

der neunten Jahrgangsstufe: die Zielkoordination und der diagnostisch-methodische Austausch. Die Lehrkräfte verständigen sich an erwartungswidrig schwachen Schulen häufiger untereinander über Leistungsanforderungen. Sie beraten ferner häufiger die Stärken und Schwächen der Schülerinnen und Schüler sowie ihre eigene Unterrichtspraxis und beruflichen Bewältigungsstrategien. In Schulen, in denen die Schülerinnen und Schüler unterdurchschnittliche Testleistungen erzielen, reagieren die Fachkollegien offenbar durch verstärkte Kooperation auf diese Problemlagen. Dieses Zusammenhangsmuster zeigt sich mit den gleichen Effektgrößen nicht nur bei den sprachlichen Kompetenzen im Deutschen insgesamt, sondern auch bei der Sprachbewusstheit und der Textproduktion (Semantik/Pragmatik) – Kompetenzen, bei denen die Schülerinnen und Schüler im Deutschen signifikante Zuwächse erzielen.

Nach Abbildung 34.10 weisen Schülerinnen und Schüler an Schulen, an denen die Fachkollegien Englisch – wiederum ohne Kenntnis der Testergebnisse – die Einschätzung teilen, dass ihre Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe die in DESI gestellten Anforderungen an die Sprachkompetenzen im Englischen mehrheitlich erfüllen und beispielsweise einem Austauschschüler einen Brief über ihre Schule und Freizeit schreiben können, auch höhere Englischkompetenzen auf ($\beta = .14$). In den Fachkollegien Englisch besteht zwar ein engerer Zusammenhang zwischen den Leistungseinschätzungen und der Fokussierung auf die einzelnen sprachlichen Kompetenzen im Unterricht als in den Fachkollegien Deutsch, doch hat auch hier die Fokussierung auf Sprachkompetenzen im Englischunterricht keinen signifikanten Effekt auf die Schülerkompetenzen. Zwischen den Leitvorstellungen und den kooperativen Praktiken besteht auch bei den Fachkollegien Englisch kein Zusammenhang. In Schulen mit erwartungswidrig niedrigen Englischkompetenzen arbeiten die Lehrkräfte des Fachkollegiums Englisch jedoch stärker zusammen ($\beta = -.08$). Offenbar verständigen sich die Fachkollegien verstärkt über sprachliche Aufgabenstellungen, Leistungsanforderungen und Beurteilungsmaßstäbe, wenn das Kompetenzniveau den Erwartungen nicht entspricht. Kooperation wird auch im Fachkollegium Englisch eher reaktiv als proaktiv praktiziert. Diese Zusammenhänge zeigen sich mit ähnlichen Effektgrößen nicht nur bei den Englischkompetenzen insgesamt, sondern auch bei der Textrekonstruktion und dem Hörverstehen – Kompetenzen, bei denen die Schülerinnen und Schüler im Englischen signifikante Zuwächse erzielen.

Insgesamt gehen die sprachlichen Kompetenzen der Schülerinnen und Schüler im Deutschen und im Englischen mit höheren Leistungseinschätzungen in den Fachkollegien einher. Der fehlende Zusammenhang zwischen den kompetenzbezogenen Leitvorstellungen und der Zusammenarbeit in den Fachkollegien ist möglicherweise ein Grund dafür, weshalb die Fokussierung auf einzelne sprachliche Kompetenzen und die Kooperationspraktiken keinen Effekt auf das kumulative Kompetenzniveau haben.



Modellfit: $\chi^2 = 160.407$; $df = 79$; CFI = .96; RMSEA = .01; $SRMR_{\text{between}} = .07$; $R^2_{\text{between}} = .87$.

Abbildung 34.10: Zusammenhang zwischen den professionellen Leitvorstellungen und kooperativen Praktiken im Fachkollegium Englisch und den Schülerkompetenzen im Englischen unter Berücksichtigung der in Tabelle 34.3 aufgeführten Kontrollvariablen.

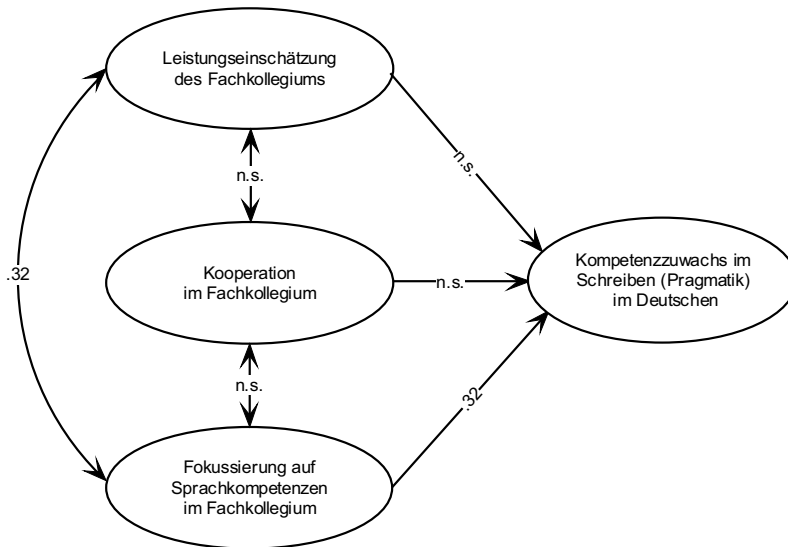
Als Quintessenz dieser Analysen kann festgehalten werden, dass die Leitvorstellungen und kooperativen Praktiken in den Fachkollegien kaum aufeinander abgestimmt sind. Ob eine höhere Interdependenz zwischen professionellen Leitvorstellungen und Praktiken hinsichtlich der sprachlichen Kompetenzen effektiv wäre, ließe sich allerdings nur in einem Interventionsdesign oder zumindest in einer vollständigen Längsschnittuntersuchung überprüfen.

34.5 Zusammenhänge der professionellen Leitvorstellungen und kooperativen Praktiken in den Fachkollegien und mit dem Kompetenzzuwachs der Schülerinnen und Schüler

Da in DESI ausgewählte Kompetenzen jeweils am Anfang und am Ende der neunten Jahrgangsstufe erhoben wurden, lässt sich überprüfen, ob die Leitvorstellungen und kooperativen Praktiken der Fachkollegien auf Schulebene einen Effekt auf die Kompetenzzuwächse der Schülerinnen und Schüler innerhalb eines Jahres haben. Die Verwendung von Zuwachs- bzw. Differenzwerten der Sprachkompetenzen als Maß für den mit der Schule bzw. mit den institutionellen Praktiken der Fachkollegien auf Schulebene verbundenen *value added* im Hinblick auf die Kompetenzen wird damit begründet, dass Zuwachs- bzw. Differenzwerte unverfälschte erwartungstreue Schätzer für die wahre Veränderungen der Schülerkompetenzen darstel-

len (vgl. Kap. 3; Nachtigall/Suhl 2002). Analog zu den vorangegangenen Analysen wurden in Mehrebenenstrukturgleichungsmodellen die um die individuellen Schülermerkmale und um die Zusammensetzung der Schülerschaft bereinigten Effekte der Schulmerkmale auf die Kompetenzzuwächse der Schülerinnen und Schüler geschätzt. Als Prädiktoren wurden die im vorangegangenen Abschnitt beschriebenen Dimensionen *Leistungseinschätzungen*, *Kooperation im Fachkollegium* und *Fokussierung auf Sprachkompetenzen* verwendet.

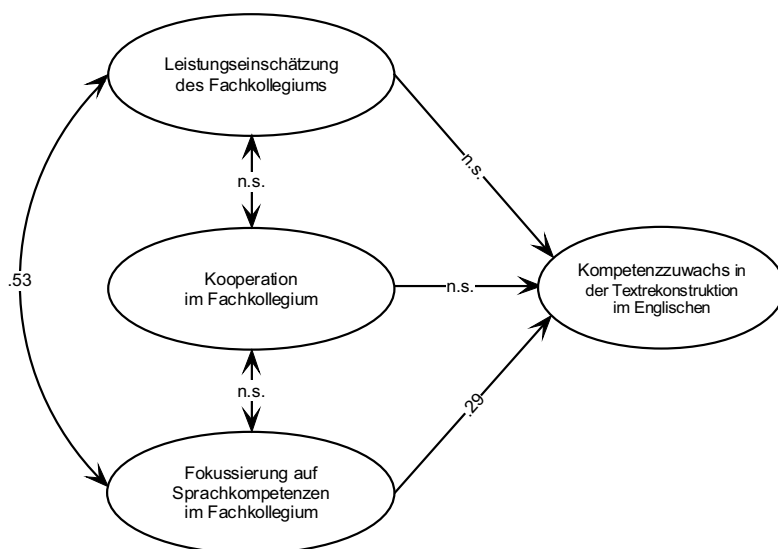
Die Zusammenhänge zwischen den Leitvorstellungen und kooperativen Praktiken der Fachkollegien und den Kompetenzzuwächsen unterscheiden sich deutlich von den Zusammenhängen für den kumulativen Leistungsstand der Schülerinnen und Schüler (vgl. Abbildung 34.9 und 34.11 versus 34.12 und 34.13). Die Höhe der Leistungserwartungen und das Ausmaß der Kooperation in den Fachkollegien haben jeweils keinen signifikanten Effekt auf die Kompetenzzuwächse der Schülerinnen und Schüler im Deutschen und im Englischen. In Schulen mit hohen Leistungserwartungen werden weder im Deutschen noch im Englischen überdurchschnittliche Zuwächse erzielt. Wenn die Fachkollegien umgekehrt die Schülerkompetenzen eher niedrig einschätzen, zeichnet sich auch keine unterdurchschnittliche Leistungsentwicklung ab. Ob hier möglicherweise Leistungspotentiale unausgeschöpft bleiben, kann aufgrund des partiellen Längsschnittdesigns nicht beantwortet werden. Ebenso trägt die Kooperation im Fachkollegium, die verstärkt in Schulen mit niedrigem Kompetenzniveau praktiziert wird, nicht zu einer unterdurchschnittlichen Kompetenzentwicklung bei, ist aber offenbar auch nicht so zielgerichtet, dass überproportionale Leistungszuwächse erzielt werden. Allerdings hat die gemeinsame Fokussierung der Fachkollegien auf die sprachlichen Kompetenzen im Unterricht einen positiven Effekt auf den Kompetenzzuwachs der Schülerinnen und Schüler im Deutschen ($\beta = .32$) und im Englischen ($\beta = .29$). Ein höherer Kompetenzzuwachs zeichnet sich dann ab, wenn die Fachkollegien die einzelnen sprachlichen Kompetenzen als Schwerpunkte im Unterricht als wichtig einschätzen. In beiden Fächern messen die Fachkollegien der mündlichen Kommunikationsfähigkeit zwar das höchste Gewicht zu, aber zur Varianz zwischen Schulen trägt vor allem die unterschiedlich starke Berücksichtigung der schriftlichen Kommunikationsfähigkeit, des korrekten Sprachgebrauchs sowie (im Englischen) des Leseverstehens bei.



Modellfit: $\chi^2 = 151.044$; $df = 80$; CFI = .92; RMSEA = .01; SRMR_{between} = .07; $R^2_{\text{between}} = .21$.

Abbildung 34.11: Zusammenhang zwischen den professionellen Leitvorstellungen und kooperativen Praktiken im Fachkollegium Deutsch und dem Kompetenzzuwachs im Schreiben (Pragmatik) im Deutschen Berücksichtigung der in Tabelle 34.3 aufgeführten Kontrollvariablen.

Einen statistisch signifikanten Zusammenhang mit dem Leistungszuwachs innerhalb der neunten Jahrgangsstufe zeigen die aggregierten Urteile der Fachkollegien in den Kompetenzdimensionen, die Veränderungen in produktiven Sprachkompetenzen indizieren: Schreiben bzw. Textproduktion (Semantik/Pragmatik) im Deutschen sowie Textrekonstruktion (C-Test) im Englischen. Diese Effekte auf den Kompetenzzuwachs sind deutlich stärker als die entsprechenden Effekte auf das Kompetenzniveau und durchaus substanziell. Für die Sprachbewusstheit im Deutschen und das Hörverstehen im Englischen – Dimensionen, in denen die Schülerinnen und Schüler auch Leistungszuwächse erzielen – zeigen sich diese Zusammenhänge nicht. Dazu müsste vermutlich das Konstrukt der Fokussierung auf Sprachkompetenzen im Hinblick auf die Beherrschung sprachlicher Normen und die Verstehensleistungen in kommunikativen Kontexten differenzierter operationalisiert werden. Gleichwohl lässt sich aus dem Zusammenhang zwischen der Fokussierung auf sprachliche Kompetenzen in den Fachkollegien und dem Kompetenzzuwachs der Schülerinnen und Schüler bei den produktiven Sprachkompetenzen schließen, dass ein gemeinsames Ziel- und Aufgabenverständnis in den Fachkollegien Deutsch und Englisch als Ansatzpunkt für konsistentes und professionelles Unterrichtshandeln genutzt werden und letztlich der sprachlichen Kompetenzentwicklung der Schülerinnen und Schüler dienen kann.



Modellfit: $\chi^2 = 156.988$; $df = 79$; CFI = .92; RMSEA = .01; SRMR_{between} = .06; $R^2_{\text{between}} = .15$.

Abbildung 34.12: Zusammenhang zwischen den professionellen Leitvorstellungen und kooperativen Praktiken im Fachkollegium Englisch und dem Kompetenzzuwachs in der Textrekonstruktion (C-Test) im Englischen unter Berücksichtigung der in Tabelle 34.3 aufgeführten Kontrollvariablen.

Analog zu den Befunden der Analysen zur Wirksamkeit des Deutsch- und Englischunterrichts in den Kapiteln 28 und 32 zeigt sich bei der Untersuchung der Leitvorstellungen und Praktiken der Fachkollegien, dass ein gezielter Fokus auf sprachliche Fähigkeiten und klare, anspruchsvolle Zielsetzungen für die sprachlichen Kompetenzen bedeutsam sind. Eine engere und prospektive Abstimmung der Leitvorstellungen und der kooperativen Praktiken in den Fachkollegien sowie eine gemeinsame Fokussierung auf die einzelnen Sprachkompetenzen könnten die Wirksamkeit unterrichtlichen Handelns auf Schulebene vermutlich steigern. Eine systematische Analyse der Wirkungszusammenhänge und Wirkungsrichtungen setzte allerdings eine Erfassung der Schul- und Leistungsmerkmale im Längsschnitt oder im Rahmen einer Interventionsstudie voraus (vgl. Creemers/Kyriakides 2006), die im Rahmen von DESI nicht realisiert werden konnte. Wenn pädagogische Maßnahmen und Aktivitäten der Fachkollegien und Schulleitung gebündelt und systematisch umgesetzt werden, könnte dies mittelfristig zu nachhaltigen Effekten führen. Für den mathematischen und naturwissenschaftlichen Unterricht hat das Modellprogramm „SINUS“ eindrucksvoll belegt, dass gezielte und kooperative Aktivitäten zur Verbesserung des Unterrichts auf Schulebene zu einer Steigerung der Schülerkompetenzen führen können (Prenzel u.a. 2005).

34.6 Institutionelle Bedingungen des Sprachunterrichts aus der Sicht der Schulleitungen

Nachdem in den vorigen Abschnitten die professionellen Leitvorstellungen und Praktiken der Fachkollegien im Mittelpunkt standen, werden im Folgenden die schulischen Angebote und Praktiken aus der Sicht der Schulleitungen behandelt. Die Schulleitungen sind neben der Vertretung der Schule nach Außen vor allem für eine effiziente und effektive Organisation des Schul- und Unterrichtsbetriebs verantwortlich. Durch ihre koordinierende Funktion in Fragen der Organisation, des Personals und des Unterrichts haben sie eine besondere Bedeutung für das Profil und die Gestaltung einer Schule, die in den besonderen Angeboten, Aktivitäten und Gepflogenheiten der Schule zum Ausdruck kommen. In der DESI-Befragung der Schulleitungen wurden vor allem Angebote und Praktiken auf Schulebene erhoben, die in der Schuleffektivitätsforschung mit höheren Schülerleistungen in Zusammenhang gebracht werden (vgl. auch Abschnitt 34.3).

In DESI unterscheiden sich die Einschätzungen der Schulleitungen bezüglich des eigenen Leitungsverhaltens (pädagogische Führung), der kooperativen Aktivitäten im Gesamtkollegium (Konsens und Kooperation) und des Schulklimas (geordnete Lernumgebung, Schüler-Lehrer-Beziehungen) in einigen Aspekten zwischen den Schularten. Effekte des Schulleitungshandelns und fachunspezifischer Praktiken auf Schulebene auf die Schülerleistungen zeigen sich in DESI nicht. Dieses Ergebnis deckt sich mit Befunden aus anderen Studien (vgl. Hallinger/Heck 1996). Zu berücksichtigen ist dabei, dass die globalen Einschätzungen der Schulleitungen vor allem über das Selbstbild der Schule Auskunft geben, aber die Realität des Fachunterrichts und die fachbezogenen Schülerleistungen aus einer gewissen Distanz beurteilen. Da direkten Selbstbewertungen des eigenen Verhaltens und des Klimas an der eigenen Schule keine systematischen Vergleiche mit anderen Schulen zugrunde liegen können, gelten diese Indikatoren des Schulgeschehens als weniger valide und zuverlässig (Baumert u.a. 2004; Clausen 2002).

In DESI wurden die Schulleitungen auch nach den *spezifischen* Angeboten und Praktiken der Schule im Bereich der Sprachen, der Evaluation und Qualitätsentwicklung sowie der Elternbeteiligung befragt (vgl. Tabelle 34.4). Die Angaben zu den zusätzlichen Angeboten und Kursen im Bereich der Sprachen dienen als Indikator für kompensatorische bzw. remediale Lerngelegenheiten, die eine Schule für Schülerinnen und Schüler mit nicht deutscher Erstsprache und mit besonderem Förderbedarf bereitstellt. Die Fragen nach der Teilnahme und erzielten Preisen bei sprachlichen Wettbewerben erfassen die kompetitiven Angebote, die eine Schule für Leistungssteigerungen nutzt. Förderkurse und Wettbewerbe stellen adaptive Strategien zur Verbesserung der sprachlichen Leistungen für Schülerinnen und Schüler mit unterschiedlichen Voraussetzungen dar.

Tabelle 34.4: Leistungsorientierte Angebote und Praktiken der Schule im Bereich der Sprachen aus Sicht der Schulleitung (Skalen, Beispielitems, Mittelwert (MW), Standardabweichung (SD), Anzahl der Items und interne Konsistenz (α)).

Skalen und Beispielitems	MW	SD	Items	α
Alle Items haben ein vierstufiges Antwortformat von 1-4. Items mit * haben ein zweistufiges Antwortformat: 0, 1.				
Leistungsorientierte Angebote und Praktiken im Bereich der Sprachen				
Förderangebote für nicht deutschsprachige Schüler/innen	2.13	0.99	2	.68
<i>Förderangebote in Deutsch für Schüler/innen mit nicht deutscher Erstsprache</i>				
Profil im Bereich der Betreuungs- und Förderangebote	0.41	0.34	5	.74
<i>Betreuungsangebote am Nachmittag*</i>				
Profil in der sprachlichen Förderung von Schüler/innen mit Migrationshintergrund	0.34	0.38	2	.61
<i>Förderunterricht in Deutsch für Schüler/innen nicht deutscher Erstsprache*</i>				
Teilnahme und Preise bei Deutschwettbewerben	1.70	0.67	2	.74
<i>Teilnahme an Wettbewerben im Fach Deutsch</i>				
Teilnahme und Preise bei Englischwettbewerben	1.38	0.62	4	.88
<i>Teilnahme an Wettbewerben im Fach Englisch</i>				
Evaluation und Qualitätsentwicklung				
Evaluation für die schulinterne Qualitätsentwicklung	2.25	0.45	11	.81
<i>Die Überprüfung der Schülerleistungen nutzen wir, um den Unterricht auf die gezeigten Stärken und Schwächen der Schüler/innen einzustellen.</i>				
Evaluation des Deutsch- und Englischunterrichts durch Schülertests	0.67	0.47	2	.98
<i>Leistungsbewertungen der Schüler/innen durch Tests oder Prüfungen.*</i>				
Vergleichende Beobachtung der Schülerleistungen innerhalb der Schule	0.52	0.35	3	.62
<i>Leistungsfeststellungen an der Schule werden genutzt, um die Entwicklung des Leistungsniveaus der Schule von Jahr zu Jahr zu beobachten.*</i>				
Kriterienorientierte Diagnostik	1.79	0.41	4	.61
<i>Zur Leistungsfeststellung werden an der Schule Leistungsstandards genutzt, die in zentralen Fächern von den Schüler/innen dieser Schule erreicht werden sollen.</i>				
Elternarbeit				
Unterstützung schulischer Ziele durch die Eltern	2.83	0.56	2	.74
<i>Die Elternarbeit an unserer Schule hat dazu geführt, dass Eltern die Leistungserwartungen der Schule stärker unterstützen.</i>				
Elternbeteiligung	3.64	0.45	3	.78
<i>Besuch von Elternabenden/Elternsprechtagen</i>				
Unterstützung des Kindes durch die Eltern	2.80	0.83	2	.61
<i>Durchsehen von Hausaufgaben</i>				
Unterstützung von Förderangeboten durch die Eltern	1.63	0.54	4	.70
<i>Mitarbeit bei der Hausaufgabenbetreuung</i>				
Unterstützung außerunterrichtlicher Aktivitäten	2.29	0.60	7	.78
<i>Mitarbeit bei der Kooperation mit externen Einrichtungen</i>				

Die Erfassung der Praktiken der Leistungsfeststellung betrifft die Strategien, die eine Schule für eine evidenzbasierte Schul- und Unterrichtsentwicklung einsetzt. Die Fragen nach der Beteiligung und Mitwirkung der Eltern in der Schule dienen der Erfassung der externen Unterstützung schulischer Ziele und Angebote durch das Elternhaus.

Hinsichtlich der adaptiven Angebote der Sprachförderung zeigen sich erwartungsgemäß deutliche Schulartunterschiede. Zusätzliche Angebote im Bereich der sprachlichen Förderung von Schülerinnen und Schülern mit nicht deutscher Erstsprache finden sich vor allem in den Hauptschulen, die besondere Anstrengungen unternehmen, um auf den spezifischen Bedarf vieler ihrer Schülerinnen und Schüler einzugehen (vgl. Abbildung 34.13). Auch die Integrierten Gesamtschulen weisen ein Profil mit überdurchschnittlicher sprachlicher Förderung von Schülerinnen und Schülern mit nicht deutscher Erstsprache auf. Allerdings stellen die Gesamtschulen allgemeine Betreuungs- und Förderangebote wie z.B. Hausaufgabenhilfe oder Betreuungsangebote am Nachmittag, die es an Hauptschulen kaum gibt, häufiger bereit als spezielle Förderangebote für Schülerinnen und Schüler mit nicht deutscher Erstsprache. Auch die Gymnasien und die Realschulen profilieren sich im Bereich der allgemeinen Betreuungs- und Förderangebote, ohne allerdings einen besonderen Akzent auf die Förderung von Schülerinnen und Schüler mit nicht deutscher Erstsprache zu setzen. Die Unterschiede zwischen den Schularten in den Förderangeboten sind zwar klein, aber bedeutsam.

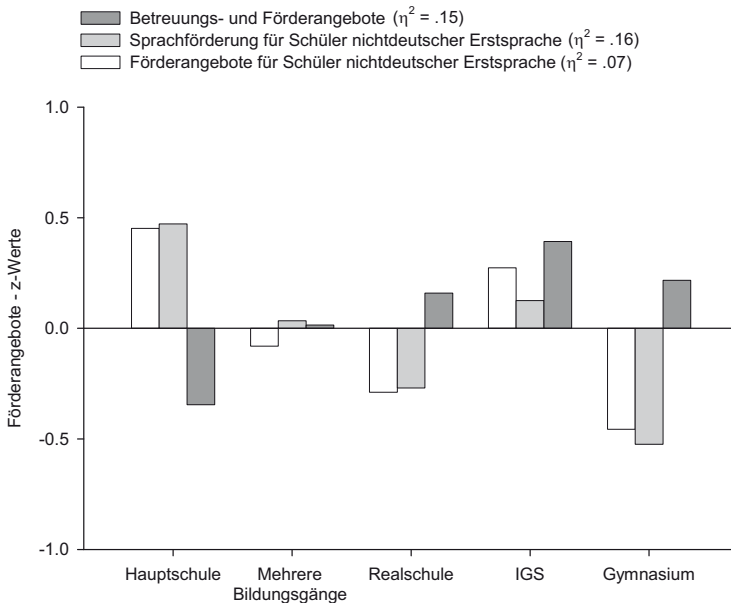


Abbildung 34.13: Zusätzliche Förderangebote in der Schule nach Schularten.

Auch bei der Teilnahme an Wettbewerben im Bereich der Sprachen zeigen sich unterschiedliche Schulartprofile (vgl. Abbildung 34.14). Die Schulartunterschiede im

Bereich der Englischwettbewerbe sind sehr groß. Dabei setzen sich die Gymnasien von allen übrigen Schularten deutlich ab. Angebote zur Förderung von Spitzenleistungen in den Fremdsprachen sind ganz überwiegend eine gymnasiale Praxis. Sie finden sich selten an Integrierten Gesamtschulen, Schulen mit mehreren Bildungsgängen und Realschulen und so gut wie gar nicht an Hauptschulen. Im Bezug auf die Teilnahme an Wettbewerben im Deutschen überschneiden sich die Häufigkeiten in den einzelnen Schularten relativ stark, so dass sich keine bedeutsamen Schulartunterschiede nachweisen lassen. Im Fach Deutsch wurden offenbar Wettbewerbsformen und –modalitäten entwickelt, die auch an die Voraussetzungen von Schülerinnen und Schüler nicht gymnasialer Bildungswege anknüpfen und nach Auskunft der Schulleitungen von den Schulen auch aufgegriffen werden.

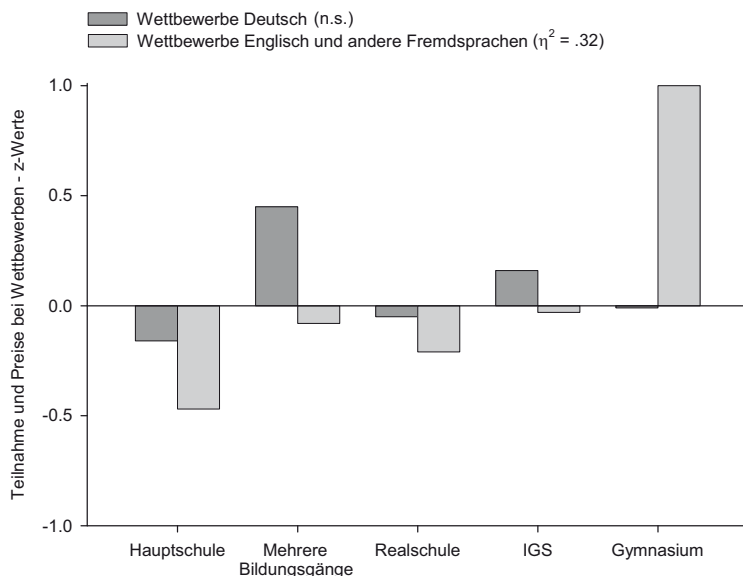


Abbildung 34.14: Teilnahme und Preise der Schule bei Wettbewerben im Bereich der Sprachen nach Schularten.

Bei den Verfahren der Evaluation von Schule und Unterricht gibt es im Hinblick auf den Einsatz von Schülertests und von Vergleichen von Schulleistungen über die Zeit keine bedeutsamen Unterschiede zwischen den Schularten (vgl. Abbildung 34.15). Die Schulleitungen von Hauptschulen berichten allerdings von einer stärkeren Orientierung an Leistungsstandards und klassenübergreifenden Leistungstests als dies bei Realschulen und Gymnasien der Fall ist. Dies dürfte vor allem Ausdruck eines erhöhten Bedarfs nach Vergewisserung über den Leistungsstand von Hauptschülerinnen und -schülern sein, damit diese die Schule nach der neunten Jahrgangsstufe erfolgreich abschließen und eine berufliche Ausbildung anschließen können. Maßnahmen der internen Evaluation und Qualitätsentwicklung werden überdurchschnittlich häufig aus den Schulen mit mehreren Bildungsgängen und den Integrierten Gesamtschulen be-

richtet, die aufgrund der Fachleistungs- bzw. Bildungsgangdifferenzierung in besonderem Maße zum Vergleich von Schülerleistungen herausgefordert werden.

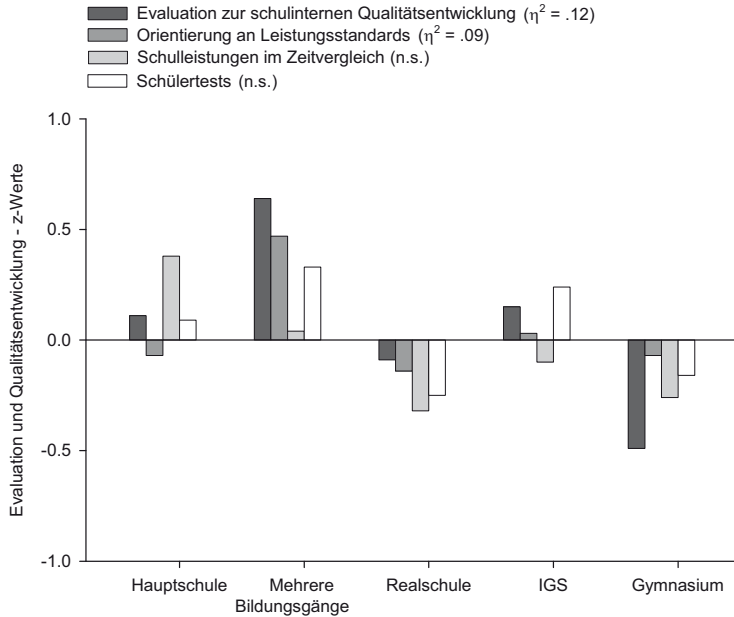


Abbildung 34.15: Maßnahmen der Evaluation und Qualitätsentwicklung in der Schule nach Schularten.

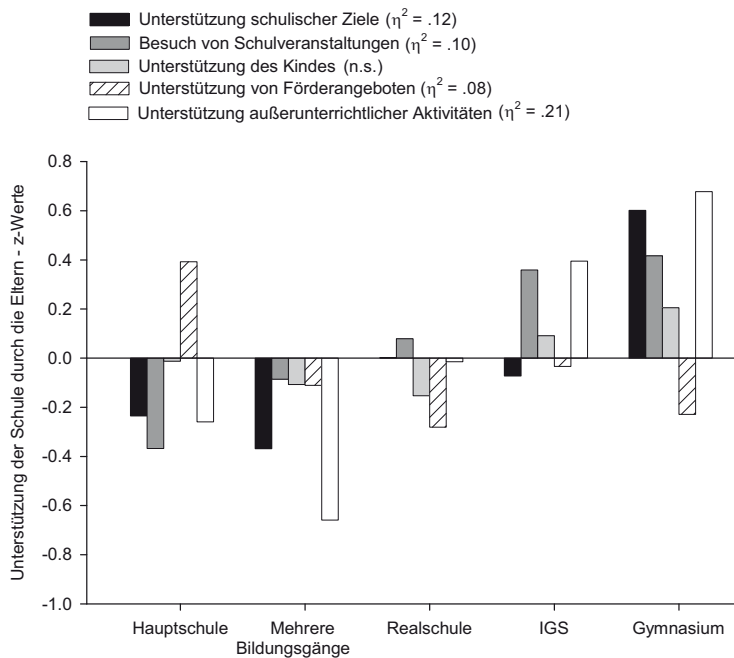


Abbildung 34.16: Unterstützung der Schule durch die Eltern nach Schularten.

Von der externen Unterstützung der Schule durch die Eltern profitieren die einzelnen Schularten in unterschiedlichem Maße (vgl. Abbildung 34.16). Die Schulartunterschiede sind, mit Ausnahme der Skala „Unterstützung des eigenen Kindes“, durchaus bedeutsam. Die Schulen mit mehreren Bildungsgängen können am wenigsten auf eine Unterstützung schulischer Ziele und eine Mitwirkung bei außerunterrichtlichen Aktivitäten durch die Eltern zählen. Ähnliches zeigt sich bei den Hauptschulen, die Schülerinnen und Schüler beschulen, die auf Unterstützung besonders angewiesen sind. Wenn die Schulleitungen angeben, dass die Eltern Förderangebote durch Beteiligung unterstützen, dürfte darin vor allem eine starke Nachfrage der Eltern nach schulischen Förderangeboten wie z.B. Hausaufgabenbetreuung zum Ausdruck kommen. Diese ist in den Realschulen und den Gymnasien unterdurchschnittlich ausgeprägt. Ansonsten ist die elterliche Unterstützung in den Gymnasien bei den außerunterrichtlichen Aktivitäten, bei der Unterstützung schulischer Ziele, die sich auf Leistungsanforderungen und das Schulklima richten, sowie beim Besuch von Schulveranstaltungen wie z.B. Elternabenden, Elternsprechtagen und Schulgremien überdurchschnittlich ausgeprägt. Die Integrierten Gesamtschulen kommen diesem Profil elterlicher Unterstützung am nächsten, jedoch mit einem Unterschied: die Schulleitungen schätzen die Unterstützung schulischer Ziele durch die Eltern nicht allzu hoch ein. Insgesamt profitieren einzelne Schularten wie die Gymnasien, aber auch die Integrierten Gesamtschulen mehr von der externen Unterstützung durch die Eltern als andere. Die hohe Akzeptanz von Förderangeboten bei Eltern in Hauptschulen könnte allerdings ein Hinweis darauf sein, dass auch Schulen, die insgesamt eher wenig auf elterliche Unterstützung zählen können, mit spezifischen Förderangeboten das Interesse von Eltern wecken und für die Elternarbeit nutzen könnten.

34.7 Zusammenhänge der institutionellen Bedingungen des Sprachunterrichts und mit den Sprachkompetenzen der Schülerinnen und Schüler

Die sprachlichen Angebote und Praktiken der Schulen zur Förderung der Sprachkompetenzen einschließlich der Beteiligung und Unterstützung durch die Eltern lassen teilweise bedeutsame Unterschiede zwischen den einzelnen Schularten erkennen. Bei den einzelnen Schulmerkmalen gibt es aber auch Überschneidungen zwischen den Schularten, die mit einer Variation dieser Praktiken zwischen den Schulen innerhalb der einzelnen Schularten einhergehen. Deshalb wurden wiederum auf der Basis von Mehrebenenstrukturgleichungsmodellen die institutionellen Angebote und Praktiken unter Verwendung der in Abschnitt 34.4 vorgestellten Kontrollvariablen auf ihre Prädiktionskraft für die sprachlichen Kompetenzen der Schülerinnen und Schüler überprüft (vgl. Tabelle 34.5).

Hierzu wurden die auf Schulleitungsangaben basierenden Skalen zu den institutionellen Bedingungen des Sprachunterrichts, analog dem Vorgehen bei den Fachkollegien, soweit möglich zu übergeordneten Faktoren zusammengefasst. Aus den in Tabelle

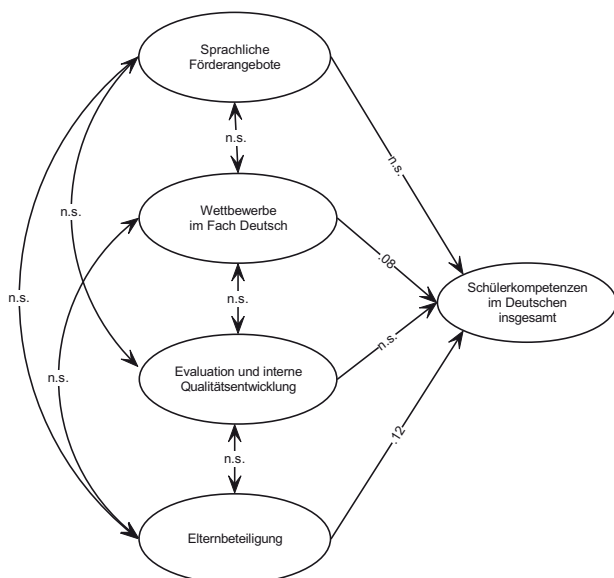
34.4 angeführten Skalen wurden – wie dort bereits gruppiert – die vier Dimensionen gebildet: die Dimension *Sprachliche Förderangebote* wird durch drei Indikatoren abgedeckt, die Dimension *Wettbewerbe im Fach Deutsch bzw. Englisch* durch je einen Indikator, die Dimension *Evaluation und Qualitätssicherung* durch vier Indikatoren und die Dimension *Elternbeteiligung* durch fünf Indikatoren.

Tabelle 34.5: Modellierung des Zusammenhangs zwischen den institutionellen Bedingungen des Sprachunterrichts und den Schülerkompetenzen im Deutschen und im Englischen.

Kontrollvariablen und Prädiktoren	Kriterien
Individualebene <ul style="list-style-type: none"> - Geschlecht - Kognitive Grundfähigkeiten - Sozioökonomischer Status: HISEI - Erstsprache - Teilnahme am bilingualen Unterricht 	<ul style="list-style-type: none"> - Gesamtleistung: Deutsch - Gesamtleistung: Englisch - Zuwachs Deutsch: Schreiben; Sprachbewusstheit - Zuwachs Englisch: Textrekonstruktion; Hörverstehen
Schulebene <ul style="list-style-type: none"> - Schularten - Mädchenanteil - Kognitive, soziale und sprachliche Zusammensetzung der Schülerschaft - Zugehörigkeit zu einer bilingualen Schule - <i>Sprachliche Förderangebote</i> - <i>Wettbewerbe im Deutschen / Englischen</i> - <i>Evaluation und Qualitätsentwicklung</i> - <i>Elternbeteiligung</i> 	

Als Ergebnis der Mehrebenenstrukturgleichungsmodelle lässt sich festhalten, dass sich die in der Schuleffektivitätsforschung identifizierten Prozessmerkmale Evaluation und Elternbeteiligung als latente Faktoren replizieren lassen. Wenn die Angebote zur sprachlichen Förderung und zur Teilnahme an Sprachwettbewerben eine gemeinsame Dimension darstellten, könnten sie als adaptive Strategie der Leistungsförderung bzw. des schulischen Curriculums im Umgang mit unterschiedlichen Lernvoraussetzungen interpretiert werden (vgl. Scheerens/Bosker 1997). In den Schulen der DESI-Stichprobe stellen diese Angebote jedoch zwei eigenständige Faktoren dar, insofern als die Schulen remediale und kompetitive Angebote und Praktiken unabhängig voneinander einsetzen. Der Befund, dass zwischen den sprachlichen Förderangeboten und Wettbewerben einerseits und den Praktiken der Evaluation und Qualitätsentwicklung andererseits nicht der geringste Zusammenhang besteht, ist ein Hinweis darauf, dass vielfältige Aktivitäten der Leistungsförderung, Evaluation und Qualitätsentwicklung unternommen, aber wenig aufeinander abgestimmt werden. Da Evaluationsverfahren auch dem Zweck dienen sollten, die Passung zwischen schulischen Angeboten und Praktiken und den Voraussetzungen, der Nutzung und den Leistungen der Schülerinnen und Schüler zu überprüfen, deuten die fehlenden Zusammenhänge zwischen diesen Maßnahmen darauf hin, dass re-

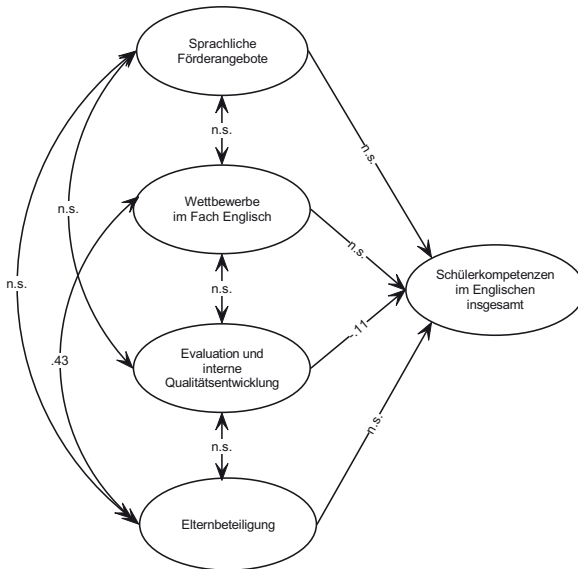
mediale und kompetitive Angebote nicht systematisch von Evaluationsmaßnahmen begleitet werden.



Modellfit: $\chi^2 = 247.670$; $df = 160$; CFI = .95; RMSEA = .01; SRMR_{between} = .09; R^2_{between} = .87.

Abbildung 34.17: Zusammenhang zwischen den institutionellen Bedingungen der Schulen und den Schülerkompetenzen im Deutschen unter Berücksichtigung der in Tabelle 34.5 aufgeführten Kontrollvariablen.

Für die Analyse der Zusammenhänge mit den Sprachkompetenzen der Schülerinnen und Schüler wurden die vier Dimensionen der institutionellen Bedingungen des Sprachunterrichts in den Schulen in Mehrebenenstrukturgleichungsmodellen als Prädiktoren für die sprachlichen Kompetenzen der Schüler auf Schulebene verwendet, als abhängige Variablen wieder die Plausible Values. In Abbildung 34.17 und 34.18 sind aus Gründen der Übersichtlichkeit nur die standardisierten, um die individuellen und institutionellen Ausgangslagen bereinigten Netto-Effekte der Konstrukte *Sprachliche Förderangebote*, *Wettbewerbe im Fach Deutsch bzw. Englisch*, *Evaluation und Qualitätsentwicklung* sowie *Elternbeteiligung* auf die Schülerkompetenzen im Deutschen und im Englischen als Ergebnisse der Zweiebenenstrukturgleichungsmodelle dargestellt. Bei der Betrachtung der Zusammenhangsmuster mit den Schülerkompetenzen zeigt sich zunächst, dass die in der Schuleffektivitätsforschung als wirksam identifizierten Dimensionen nicht in gleichem Maße für beide Kompetenzbereiche bedeutsam sind, da sich die Zusammenhänge mit den Leistungen in den Fächern Deutsch und Englisch unterscheiden. An Schulen, die stärker an Wettbewerben im Fach Deutsch teilnehmen, zeigen die Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe – unter Berücksichtigung individueller Ausgangslagen und der Zusammensetzung der Schülerschaft – höhere Leistungen im Deutschen ($\beta = .08$).



Modellfit: $\chi^2 = 268.309$; $df = 160$; CFI = .92; RMSEA = .01; SRMR_{between} = .09; $R^2_{\text{between}} = .88$.

Abbildung 34.18: Zusammenhang zwischen den institutionellen Bedingungen der Schulen und den Schülerkompetenzen im Englischen unter Berücksichtigung der in Tabelle 34.3 aufgeführten Kontrollvariablen.

An Schulen, die verstärkt an Englischwettbewerben teilnehmen – dies sind überwiegend Gymnasien –, zeigen die Schülerinnen und Schüler keine höheren Leistungen im Englischen, obwohl diese Schulen häufiger von den Eltern unterstützt werden ($r = .43$). Da die Teilnahme von Schulen an Deutsch- im Gegensatz zur Teilnahme an Englischwettbewerben an allen Schularten zu finden ist, werden kompetitive Praktiken möglicherweise erst dann bedeutsam, wenn sie von den Schülerinnen und Schülern in der Breite genutzt werden. Die Angebote zur sprachlichen Förderung, die sich an nicht deutschsprachige oder förderbedürftige Schülerinnen und Schüler im Deutschen richten, zeigen keinen Effekt auf die Sprachkompetenzen im Deutschen. Vermutlich entfalten remediale Angebote erst bei spezifischer Ausgestaltung und systematischer Nutzung ihre Wirksamkeit (vgl. Stanat 2005, 2006). Eine Überprüfung dieser Interpretation würde allerdings längsschnittliche Nutzungsdaten oder ein entsprechendes Interventionsdesign voraussetzen.

Bei den Verfahren der Evaluation und Qualitätsentwicklung gibt es ebenfalls fachspezifische Unterschiede in den Zusammenhängen mit den Schülerkompetenzen: Schülerinnen und Schüler, die Schulen besuchen, in denen Maßnahmen der Evaluation und Qualitätsentwicklung häufiger eingesetzt werden, weisen unter Berücksichtigung ihrer individuellen Ausgangslagen und der Schülerzusammensetzung auf Schulebene niedrigere Leistungen im Englischen ($\beta = -.11$), nicht aber im Deutschen auf. Die verstärkte Anwendung von Evaluationsverfahren und Aktivitäten zur Qualitätsentwicklung folgt offenbar dem gleichen Muster wie die Kooperation in den Fachkollegien Englisch: Sie werden eingesetzt, wenn der Leistungsstand der

Schülerinnen und Schüler den Kompetenzerwartungen nicht entspricht. Ein eher reaktiver Einsatz von Maßnahmen der Evaluation und Qualitätsentwicklung in der Schule kann auch deshalb angenommen werden, weil die fehlenden Zusammenhänge zwischen den sprachlichen Förderangeboten, Wettbewerben und der Evaluation und Qualitätsentwicklung wenig Kohärenz und Konsistenz im Einsatz schulischer Praktiken erkennen lassen.

In Schulen, in denen die Eltern stärker eingebunden sind, weisen die Schülerinnen und Schüler am Ende der neunten Jahrgangsstufe unter Berücksichtigung individueller Ausgangslagen und der Schülerzusammensetzung höhere Leistungen im Deutschen auf ($\beta = .12$). Bei den Englischleistungen zeigt sich auch ein positiver Zusammenhang mit der Elternbeteiligung, der allerdings die Schwelle zur statistischen Bedeutsamkeit verfehlt. Die Unterschiede in den Zusammenhängen dürften den Sachverhalt widerspiegeln, dass die Englischkompetenzen im Vergleich zu den Kompetenzen im Deutschen, das für die meisten Schülerinnen und Schüler in DESI die Erstsprache ist, in stärkerem Maße vom Schulunterricht abhängen als von der häuslichen Sozialisation und der Unterstützung der Eltern. Ein Zusammenhang zwischen der Zusammenarbeit mit den Eltern und den Sprachkompetenzen der Schülerinnen und Schüler im Deutschen zeigt sich im Übrigen nicht nur in DESI. In der International Reading Literacy Study der IEA aus dem Jahr 1990 zeigt sich in West-Deutschland ein positiver Effekt der Zusammenarbeit zwischen Elternhaus und Schule auf das Leseverständnis von Achtklässlern – auch wenn dort noch nicht mit Mehrebenenstrukturgleichungsmodellen gearbeitet wurde (vgl. Lehmann u.a. 1995 für die 9- und 14-Jährigen in West-Deutschland; zu den Ergebnissen im internationalen Vergleich vgl. Ross/Postlethwaite 1994; Munck/Lundberg 1994; Scheerens/Bosker 1997). Ho/Willms (1996) zeigen auf der Basis der Daten der National Educational Longitudinal Study (NELS) für Schülerinnen und Schüler der achten Jahrgangsstufe in US-amerikanischen Mittelschulen, dass die Elternbeteiligung sich vor allem auf die individuelle Lese- und Mathematikleistungen positiv auswirkt. Gleichwohl war dort die Elternbeteiligung für die Leseleistungen auch auf der Schulebene – nicht aber für die Mathematikleistungen – bedeutsam.

Bei der Analyse der Zusammenhänge zwischen den schulischen Angeboten und Praktiken und den Zuwächsen in den längsschnittlich erhobenen Kompetenzen im Deutschen und Englischen zeigt sich bei keinem der vier latenten Faktoren ein signifikanter Zusammenhang. Auf grafische Darstellungen wird daher verzichtet. Einzelne Indikatoren der Zusammenarbeit mit den Eltern weisen positive Effekte auf die Kompetenzentwicklung auf: Die Unterstützung von Förderangeboten durch die Eltern, z.B. durch Mitarbeit bei der Hausaufgabenbetreuung, geht mit größeren Kompetenzzuwächsen in der Sprachbewusstheit und im Lesen im Deutschen einher. Ein größerer Zuwachs in der Textrekonstruktion im Englischen ist dann zu verzeichnen, wenn Eltern schulische Ziele und Leistungserwartungen oder ihr Kind mit dem Durchsehen von Hausaufgaben stärker unterstützen. Eine aktive Elternarbeit der Schule zahlt sich demnach in spezifischen Bereichen aus.

Die Zusammenhänge zwischen der Wichtigkeit, die Lehrkräfte den sprachlichen Kompetenzen beimessen, und dem Kompetenzzuwachs der Schülerinnen und Schüler wurden auch in den Analysen zur Wirksamkeit des Deutsch- und Englischunterrichts dargelegt (vgl. Kap. 28 und 32). Die Parallelität der Befunde auf Unterrichts- und Schulebene bestätigt, welche Bedeutung ein *gezielter Fokus auf sprachliche Fähigkeiten und entsprechend klare, anspruchsvolle Zielsetzungen* haben können. Auch wenn die hier berichteten Effektgrößen – gemessen an den Standards der empirischen Sozialforschung – sich eher klein ausnehmen, sind sie – angesichts der Berücksichtigung individueller Voraussetzungen und der Schülerzusammensetzung – dennoch praktisch bedeutsam. In der Schul- und Unterrichtsforschung nehmen isolierte Maßnahmen und Qualitätsdimensionen, wenn man deren „Wirkung“ in Surveystudien wie DESI überprüft, kein exorbitantes Ausmaß an. Erst die Bündelung von pädagogischen Maßnahmen, Aktivitäten der Schulleitung und der Fachkollegien kann mittel- und langfristig, wenn sie systematisch und konzentriert eingesetzt werden, zu nachhaltigen und starken Effekten führen. Für den mathematischen und naturwissenschaftlichen Unterricht hat dies das Modellprogramm „SINUS“ eindrucksvoll belegt, das bundesweit zu einer Verbesserung der Unterrichtsqualität und auch von Schülerleistungen geführt hat. Analoge konzentrierte Maßnahmen zur Verbesserung von Schul- und Unterrichtsqualität im Bereich der Sprachen wären nach den Ergebnissen von DESI auch für den Unterricht in den sprachlichen Fächern viel versprechend (vgl. auch Kap. 20). Spezifische Interventionen wie z.B. bei der Förderung von Schülerinnen und Schülern mit nicht deutschsprachigem Hintergrund im Anschluss an PISA deuten darauf hin, dass systematische Förderprogramme, einen positiven Effekt auf die Sprachkompetenzen von Schülerinnen und Schülern mit nicht deutscher Erstsprache haben können, auch wenn die Nachhaltigkeit dieser Effekte bei kurzen Programmlaufzeiten nicht gesichert ist (vgl. Stanat o.J., 2005, 2006). Allerdings zeigen Klassen mit bilinguaalem Sachfachunterricht, der im Laufe der Sekundarstufe I systematisch aufgebaut wird und bei dem Ziele, Inhalte, Methoden und Materialien aufeinander abgestimmt sind, gegenüber vergleichbaren Klassen mit regulärem Englischunterricht, dass die sprachlichen Kompetenzen von Schülerinnen und Schülern im Englischen deutlich verbessert werden können (vgl. Kap. 35).

Literatur

- Baumert, J./Schümer, G. (2001): Schulformen als selektionsbedingte Lernmilieus. In: Deutsches PISA Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. (2001). Opladen: Leske + Budrich, S. 454-467.
- Baumert, J./Trautwein, U./Artelt, C. (2003): Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen: Leske+Budrich, S. 261-331.
- Baumert, J./Kunter, M./Brunner, M./Krauss, S./Blum, W./Neubrand, M. (2004): Mathematikunterricht aus der Sicht der PISA-Schülerinnen und Schüler und ihrer Lehrkräfte. In: Deutsches PISA Konsortium (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann, S. 314-354.

- Brophy, J.E./Good, T.L. (1986): Teacher behaviour and student achievement. In: Wittrock, M. C. (Hrsg.): *Handbook of research on teaching*. New York: Macmillan, S. 328-375.
- Clausen, M. (2002): Unterrichtsqualität: Eine Frage der Perspektive? Waxmann: Münster.
- Creemers, B.P./Kyriakides, L. (2006): Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. In: *School effectiveness and school improvement* 17, H. 3, S. 347-366.
- De Fraine, B./van Damme, J./van Landeghem, G./Opdenakker, M.-C. (2003): The effect of schools and classes on language achievement. In: *British educational research journal* 29, H. 6, S. 841-859.
- Deutsches PISA Konsortium (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich
- Deutsches PISA Konsortium (Hrsg.) (2003): PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen: Leske + Budrich
- Deutsches PISA Konsortium (Hrsg.) (2003): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann.
- Ditton, H. (2000): Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In: *Zeitschrift für Pädagogik. Qualität und Qualitätssicherung im Bildungsbereich*. Schule, Sozialpädagogik, Hochschule, Beiheft 41. Weinheim und Basel: Beltz, S. 73-92.
- Fend, H. (1998): Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim und München: Juventa.
- Gräsel, C./Fußangel, K./Pröbstel, C. (2006): Lehrkräfte zur Kooperation anregen – eine Aufgabe für Sisypchos? In: *Zeitschrift für Pädagogik* 52, H. 2, S. 205-219.
- Hallinger, P./Heck, R. (1996): Reassessing the principal's role in school effectiveness: A review of the empirical research. 1980-1995. In: *Educational Administration Quarterly* 32, S. 5-44.
- Ho, S.-C. E./Willms, J. D. (1996): Effects of parental involvement on eighth-grade achievement. In: *Sociology of Education* 69, S.126-141.
- Hoy, W. K./Tarter, C.J./Hoy, A.W. (2006): Academic Optimism of Schools: A Force for Student Achievement. In: *American Educational Research Journal* 43, H. 3, S. 425-446.
- Kruse, S./Louis, K.S./Bryk, A. (1995): An Emerging Framework for Analyzing School-Based Professional Community. In: Louis, K.S./Kruse, S. (Hrsg.): *Professionalism and Community: Perspectives on Reforming Urban Schools*. Thousand Oaks: Corvin Press, S. 23-44.
- Lehmann, R. H./Peek, R./Pieper, I./Stritzky, R. v. (1995): Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen. Weinheim und Basel: Beltz.
- Louis, K.S./Leithwood, K. (1998): From organizational learning to professional learning communities. In: Leithwood, K./Louis, K.S. (Hrsg.): *Organizational learning in schools*. Lisse: Swets en Zeitlinger, S. 275-285.
- Mortimore, P./Sammons, P./Stoll, L./Lewis, D./Ecob, R. (1988): *School matters. The junior years*. Somerset: Open Books.
- Mortimore, P. (1994): Schuleffektivität: Ihre Herausforderung für die Zukunft. In: *Bildung und Erziehung in Europa. Zeitschrift für Pädagogik, Beiheft 32*. Weinheim: Beltz, S. 117-134.
- Munck, I./Lundberg, I. (1994): Multivariate Analysis of Data from Population A. In: Elley, W. B. (Hrsg.): *The IEA Study of Reading Literacy: Achievement and Instruction in Thirty-Two School Systems*. Exeter: Pergamon, S. 193-222.
- Muthén, L.K./Muthén, B.O. (1998-2004). *Mplus User's Guide. Statistical Analysis With Latent Variables*. Los Angeles, CA: Muthén & Muthén.
- Nachtigall, C./Suhl, U. (2002). Der Regressionseffekt. Mythos und Wirklichkeit. In: *methevalreport* 4, H. 2, S. 1-9.
- Opdenakker, M.-C./van Damme, J. (2000): Effects of schools, teaching staff, and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. In: *School effectiveness and school improvement* 11, H. 2, S. 165-196.

- Prenzel, M./Carstensen, C.H./Senkbeil, M./Ostermeier, C./Seidel, T. (2005): Wie schneiden SINUS-Schulen bei PISA ab? Ergebnisse der Evaluation eines Modellversuchsprogramms. In: Zeitschrift für Erziehungswissenschaft 8, H. 4, S. 540-562.
- Prenzel, M./Senkbeil, M./Drechsel, B. (2004): Kompetenzunterschiede zwischen Schulen. In: Deutsches PISA Konsortium (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann, S. 292-296.
- Radisch, F./Steinert, B. (2005): Schulische Rahmenbedingungen im internationalen Vergleich. In: Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Valtin, R./Walther, G. (Hrsg.): ILGU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien. Münster: Waxmann, S. 159-186.
- Raudenbusch, S./Bryk, A./Congdon, R. (2005) HLM 6. Hierarchical Linear and Nonlinear Modeling. Scientific Software International.
- Reynolds, D./Sammons, P./Stoll, L./Barber, M./Hillman, J. (1996): School Effectiveness and School Improvement in the United Kingdom. In: School Effectiveness and School Improvement 7, H. 2, S. 133-158.
- Rolff, H.-G. (1980): Soziologie der Schulreform. Weinheim und Basel: Beltz.
- Rolff, H.-G. (1991): Schulentwicklung als Entwicklung von Einzelschulen? In: Zeitschrift für Pädagogik 37, H. 6, S. 865-886.
- Ross, K. N./Postlethwaite, T. N. (1994): Differences among Countries in School Resources and Achievement. In: Elley, W. B. (Hrsg.): The IEA Study of Reading Literacy: Achievement and Instruction in Thirty-Two School Systems. Exeter: Pergamon, S. 123-147.
- Sammons, P./Thomas, S./Mortimore, P. (1997a): Forging links: Effective schools and effective departments. London: Chapman.
- Sammons, P./Thomas, S./Mortimore, P. (1997b): Differential school effectiveness. Departmental variations in GCSE attainment. In: The School Field 8, H. 1-2, S. 97-125.
- Scheerens, J./Bosker, R. (1997): The Foundations of Educational Effectiveness. Oxford: Pergamon.
- Senkbeil, M./Drechsel, B./Rolff, H.-G./Bonsen, M./Zimmer, K./Lehmann, R.H./Neumann, A. (2004): Merkmale und Wahrnehmungen von Schule und Unterricht. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster: Waxmann, S. 296-314.
- Slavin, R.E. (1996): Education for all. Lisse: Swets & Zeitlinger.
- Stanat, P. (o.J.): Förderung von deutschen Sprachkompetenzen bei Kindern zugewandeter und sozial benachteiligter Familien. Unveröffentlichtes Manuskript.
- Stanat, P. (2005): Deutschstunde: Verben pauken und Theater spielen. In Sommercamps lernen Einwandererkinder Deutsch. Forscherin Petra Stanat über eine gelungene Kombination / (Interview mit Petra Stanat). In: Die Zeit, 30. Juni 2005.
- Stanat, P. (2006): Disparitäten im schulischen Erfolg: Forschungsstand zur Rolle des Migrationshintergrunds. In: Unterrichtswissenschaft 36, H. 2, S. 98-124.
- Steinert, B./Klieme, E./Maag Merki, K./Döbrich, P./Halbheer, U./Kunz, A. (2006): Lehrerverkooperation in der Schule: Konzeption, Erfassung, Ergebnisse. In: Zeitschrift für Pädagogik 52, H. 2, S. 185-204.
- Stringfield, S. (1994): A model of elementary school effects. In: Reynolds, D./Creemers, B.P. M./Nesselrodt, P.S./Teddle, C./Shaffer, E.C./Stringfield, S. (Hrsg.): Advances in school effectiveness research and practice. Oxford: Pergamon, S. 153-187.
- Teddle, C./Reynolds, D. (2001): The international handbook of school effectiveness research. London: Palmer
- Wang, M./Haertel, G.D./Walberg, H.J. (1993): Toward a knowledge base for school learning. In: Review of Educational Research 63, H. 3, S. 249-294.

Weiß, M./Steinert, B. (2001): Institutionelle Vorgaben und ihre aktive Ausgestaltung – Die Perspektive der deutschen Schulleitungen. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich, S. 427-454.

Günter Nold / Johannes Hartig / Silke Hinz / Henning Rossa

35 Klassen mit bilingualem Sachfachunterricht: Englisch als Arbeitssprache

„Englisch bilingual“, „Englisch als Arbeitssprache“ oder „bilingualer Sachfachunterricht Englisch“ sind Begriffe, die verwendet werden, um Schulprogramme zu bezeichnen, die das Ziel haben, Englisch in nicht-sprachlichen Schulfächern als Unterrichts- oder Arbeitssprache zu verwenden. Dabei ist zu beachten, dass im nicht-sprachlichen Sachfachunterricht mit Englisch als Unterrichtssprache der Gegenstand des Sachfaches im Mittelpunkt steht, nicht Fremdsprachenlernen. Da die entsprechenden Lehr- und Lernprozesse jedoch weitgehend in englischer Sprache erfolgen, werden dennoch sprachliches, interkulturelles und sachfachspezifisches Lernen miteinander verbunden. Entsprechend wird eine erhöhte Sprachkompetenz sowie die Förderung und Vertiefung interkultureller und sachfachspezifischer Kompetenzen erwartet (z.B. Hallet 2005). Bezogen auf den Spracherwerb Englisch wird davon ausgegangen, dass die Schülerinnen und Schüler, die an diesen Programmen teilnehmen, durch den vermehrten Gebrauch der Fremdsprache in der natürlichen Kommunikation eines Sachfaches einen erheblichen sprachlichen Zugewinn erlangen.

In der Bundesrepublik Deutschland findet die Idee bilingualen Unterrichts, die vor mehr als 30 Jahren im Rahmen von deutsch-französischen Bildungsgängen entwickelt wurde, sowohl bezogen auf verschiedene Sprachen als auch unterschiedliche Schularten zunehmend Zuspruch. Mittlerweile machen über 600 Schulen bilinguale Unterrichtsangebote (vgl. Thürmann 2000, 2002; Bonnet 2004). Englisch wird als Unterrichtssprache inzwischen an über 300 Schulen in der gesamten Bundesrepublik – mit einem Schwerpunkt vor allem in Nordrhein-Westfalen, aber zunehmend auch in Ländern wie Berlin, Baden-Württemberg, Hessen und Niedersachsen – in Fächern wie Geschichte, Geographie, Politik, Biologie oder auch Kunst und Sport verwendet.

In den Bundesländern werden verschiedene Konzeptionen des bilingualen Unterrichts angewandt, wobei insbesondere in Nordrhein-Westfalen an den entsprechenden Schulen durch die Einrichtung eines sprachlichen Vorlaufs besonders günstige Bedingungen geschaffen wurden. So wird hier in den Klassen fünf und sechs der Stundenanteil des Englischunterrichts auf sechs bis sieben Stunden statt der üblichen vier bis fünf erhöht, und der Einstieg in das Sachfach wird durch eine weitere zusätzliche Unterrichtsstunde in der siebten bzw. achten Klasse erleichtert. Die bisher vorherrschende Form zielt ab auf eine fachlich begrenzte, aber möglichst durchgängige Verwendung der fremden Sprache als Unterrichtssprache von der siebten Klassenstufe an. Das Deutsche wird dabei vor allem zur Entwicklung einer fachsprachlichen Zweisprachigkeit in den Unterricht einbezogen. Eine zweite Variante

des bilingualen Unterrichts hat sich in neuerer Zeit entwickelt. Hier wird die fremde Sprache als Unterrichtssprache ohne einen sprachlichen Vorlauf für zeitlich begrenzte Unterrichtsphasen und Themenbereiche in verschiedenen Sachfächern verwendet, um einerseits Schwierigkeiten, die durch den Gebrauch der Fremdsprache bedingt sein können, für den Sachfachunterricht zu minimieren und andererseits dennoch die positiven Auswirkungen der vermehrten Verwendung der Fremdsprache erzielen zu können (vgl. Christ 2002; Krechel 2003).

Auch in Nachbarländern wie beispielsweise Österreich lässt sich eine vergleichbare Entwicklung feststellen. Im europäischen Kontext haben diese Entwicklungen inzwischen zu einer auch von der Europäischen Kommission unterstützten innovativen europäischen Spracheninitiative geführt, die als Rahmenprogramm unter der Bezeichnung CLIL (*Content and Language Integrated Learning*) die verschiedenen Entwicklungen in den europäischen Ländern zusammenfasst (vgl. Marsh 2002). Es wird erwartet, dass die Schüler und Schülerinnen durch die bilingualen Programme dazu befähigt werden, eine möglichst umfassende Zweitsprachenkompetenz zu entwickeln.

Im bisher regulären Englischunterricht der Schule in Deutschland kommt der allgemeinen fremdsprachlichen Kommunikationsfähigkeit zwar eine vorrangige Stellung zu, jedoch lässt sich durch die Beschränkung des Gebrauchs der Fremdsprache auf den Englischunterricht (im Falle von Englisch) eine umfassende sprachlich-kommunikative und pragmatisch-interkulturelle Kompetenz nur in Ausnahmefällen erreichen.

Mit der Verwirklichung eines bilingual konzipierten Unterrichts ergeben sich Veränderungen nicht nur in dem entsprechenden fremdsprachlich unterrichteten Sachfach, sondern darüber hinaus auch im Fremdsprachenunterricht selbst, da die Schüler und Schülerinnen die Fremdsprache sehr viel stärker als Kommunikationsmedium erleben und verwenden; infolgedessen lassen sich schon zu einem früheren Zeitpunkt oder in komplexerer Form kulturell-landeskundliche und interkulturelle Zielsetzungen verwirklichen, die gerade auch in den Lehrplänen der verschiedenen Länder in den Vordergrund gerückt werden. So liegt zum Beispiel ein wesentliches Merkmal des bilingualen Sachfachunterrichts darin, dass Beispiele aus zielsprachigen Kulturen ausgewählt werden, die mit eigenkulturellen Inhalten verglichen werden können und damit auch eine Außensicht auf die eigene Kultur eröffnen. Der bilinguale Sachfachunterricht ermöglicht somit eine Reflexion kulturübergreifender Phänomene, auf deren Basis ein vertieftes Verständnis der eigenen Kultur entstehen kann (vgl. Hallet 1998). Gleichzeitig fördert bilingualer Unterricht sprachliche Kompetenzen. Durch den Umgang mit umfangreichem Input authentischer Materialien lernen die Schülerinnen und Schüler auf rezeptiver Ebene, Texte sprachlich und inhaltlich zu entschlüsseln. Auf sprachproduktiver Ebene erstellen sie Textformate, die dem sachfachlichen Diskurs entsprechen (z.B. beschreiben, erklären, schlussfolgern). Der bilinguale Unterricht bahnt damit insbesondere die Entwicklung fachsprachlich angemessener bzw. fachdiskursspezifischer sprachlicher Elemente an (vgl. Vollmer 2000a, b; Zydatiś 2002). Dementsprechend steht die Entwicklung

einer *Cognitive Academic Linguistic Proficiency* (CALP), d.h. kognitiv-sprachlich komplexer Fähigkeiten und Fertigkeiten, im Vordergrund, während der Erwerb von *Basic Interpersonal Communication Skills* (BICS), d.h. von Alltagskommunikation, vornehmlich im Englischunterricht stattfindet.

Zur Förderung der CALP-Komponente sieht ein typisches CLIL-Lernarrangement die Bereitstellung von *Language Support* vor, der die sorgfältige Steuerung sprachlicher Anforderungen gewährleistet. Beispiele für derartige „*support systems*“ (Thürmann 2002, S. 82ff) sind verstärkt Visualisierungen sowie die Kombination von mutter- und fremdsprachlichen Texten, die sprachliche Lernprozesse sowohl auf rezeptiver als auch auf produktiver Ebene stützen (für eine Übersicht über verschiedene *support systems* siehe Wildhage/Otten 2003; vgl. auch Helbig 2001; Bonnet 2004).

Mit dem Konzept des bilingualen Sachfachunterrichts nimmt die deutsche Schule gesellschaftliche Erwartungen ernst, die in zunehmendem Maße im Zusammenhang mit der sich entwickelnden europäischen Integration vor allem in Hinsicht auf allgemein gesellschaftliche und berufliche Qualifikationen – vorzugsweise im kaufmännischen, aber auch technischen Bereich großer und gerade auch mittelständischer Betriebe – eingefordert werden (vgl. Finkbeiner 1995). Für die Zukunft ist mit einem weiteren Ausbau bilingualer Züge in den Schulen zu rechnen. Inzwischen wurden an Universitäten erste bilinguale Zusatzstudiengänge eingerichtet, und auch in der Referendarausbildung wird vermehrt bilingualer Unterricht als Ausbildungsziel berücksichtigt.

35.1 Englisch bilinguale Teilstichprobe

Das für die DESI-Untersuchung entwickelte Forschungs- und Testinstrumentarium wurde auf eine spezielle Teilstichprobe von ca. 1200 Schülerinnen und Schüler ausgedehnt. Unter Berücksichtigung von Klassengröße und regionaler Verteilung über die gesamte Bundesrepublik war eine Untersuchung in mindestens 45 Klassen mit Englisch als Arbeitssprache im bilingualen Sachfachunterricht geplant. Die entsprechenden Schulen wurden auf Grund der unterschiedlichen regionalen Verteilung und der vergleichsweise begrenzten Anzahl in der Bundesrepublik nach festgelegten Kriterien (regional, schulartspezifisch, zumindest zweizügige Schulen mit einem bilingualen Programm in jeweils einer Klasse, Eigenschaften des bilingualen Programms an den Schulen) sortiert und danach entsprechend dem Zufallsprinzip ausgewählt.

Bei der Testdurchführung stellte sich jedoch heraus, dass in einigen Fällen nur Teile von Klassen an einem bilingualen Programm teilnahmen. Bei der Auswertung der Daten wurden jedoch möglichst vollständige Klassensätze zu Grunde gelegt, um eine Vergleichbarkeit der Ergebnisse mit nicht-bilingualen Klassen zu ermöglichen. Die Auswertung wurde weiter dadurch eingeschränkt, dass nur Klassen berücksichtigt wurden, in denen mindestens 95% der Schülerinnen und Schüler an dem bilingualen Programm über die gesamte Zeit von der fünften bis zur neunten Klasse – in

einer Reihe von Bundesländern auch von der siebten bis zur neunten Klasse – teilgenommen hatten. Damit reduzierten sich die analysierten bilingualen Klassen auf 38 Klassen mit insgesamt 958 Schülerinnen und Schülern, bei denen zumindest in einem Sachfach, in der Mehrzahl der Fälle in zwei bis drei Fächern Englisch als Unterrichtssprache verwendet wurde. Entsprechend der Verteilung der bilingualen Klassen auf die Schularten stammten 31 Klassen aus Gymnasien, 7 aus Realschulen. Gesamtschulen kamen bei der Zufallsstichprobe auf Grund der Auswahlkriterien (vor allem vollständige Klassen) nicht zum Zuge, obgleich auch in dieser Schulart inzwischen ein beachtlicher Ausbau bilingualer Angebote zu verzeichnen ist.

35.2 Ergebnisse

Bei der Erhebung der bilingualen Teilstichprobe wurde davon ausgegangen, dass die bilingualen Schülerinnen und Schüler auf Grund bestimmter Kriterien nach einem Ausleseprozess in die Programme aufgenommen wurden. Daher wurden entsprechende Daten erfragt. Die Auswertung ergab, dass die Vergleichbarkeit der Leistungen mit nicht-bilingualen Schülerinnen und Schülern nur gewährleistet ist, wenn eine Kontrollgruppe mit vergleichbaren Hintergrundvariablen gebildet wird. Dementsprechend wurde auf der Basis der Leistungswerte Deutsch zu Beginn der neunten Jahrgangsstufe, des sozioökonomischen Status der Eltern (HISEI), der kognitiven Grundfähigkeit (KFT numerisch), des Bildungsgangs, der Erstsprache sowie des Geschlechts eine Vergleichsgruppe von 987 Schülerinnen und Schülern gebildet. Die Aussagen über den Einfluss der bilingualen Programme sind im Folgenden auf diese Vergleichsgruppe bezogen. Damit ist es möglich, die Stärke der Einflüsse des bilingualen Unterrichts auf der Basis einer angemessenen Vergleichsgruppe zu bewerten.

Die in Abbildung 35.1 dargestellten Mittelwerte in den DESI-Englishtests belegen, dass die Leistungen der Schülerinnen und Schüler in bilingualen Programmen trotz Berücksichtigung der Hintergrundvariablen durchgängig erhebliche positive Einflüsse in den verschiedenen Testelementen von DESI¹ aufweisen (vgl. Abbildung 35.1). Zusätzlich wurde der Effekt der bilingualen Programme auf Basis von Regressionsanalysen in der Gesamtstichprobe geschätzt, wobei die genannten Hintergrundvariablen als Kontrollvariablen verwendet wurden. Die in Abbildung 35.2 dargestellten Ergebnisse dieser Analysen zeigen, dass die Einflüsse des bilingualen Sachfachunterrichts auf das Hörverstehen so deutlich sind, dass sie sogar in den Zuwachswerten innerhalb der neunten Jahrgangsstufe statistisch signifikant werden.

Bezieht man die in Abbildung 35.1 dargestellten Mittelwerte auf die Kompetenzniveaus in den verschiedenen Kompetenzbereichen, ist festzustellen, dass die Schülerinnen und Schüler in bilingualen Programmen im Vergleich zu der Vergleichsgruppe in vier von sechs Bereichen das jeweils höhere Kompetenzniveau erreichen, und zwar in Hörverstehen (Kompetenzniveau C), Leseverstehen

¹ Die Werte zum Sprechen sind hier nicht berücksichtigt.

(Kompetenzniveau C), Sprachbewusstheit Grammatik (Kompetenzniveau B) und globaler Sprachstand (Kompetenzniveau C). Die bilinguale Schülergruppe zeichnet sich jedoch auch in ihrer Sprachbewusstheit sprachliches Handeln (Soziopragmatik; oberes Kompetenzniveau B) sowie in ihrer Schreibkompetenz (oberes Kompetenzniveau C) durch eine jeweils höhere Kompetenzausprägung innerhalb eines Niveaus aus. Für den Bereich des Sprechens zeichnet sich eine entsprechende Tendenz ab, auch wenn hier die Ergebnisse auf Grund der eingeschränkten Stichprobe (vgl. Kapitel 16) gesondert zu betrachten sind.

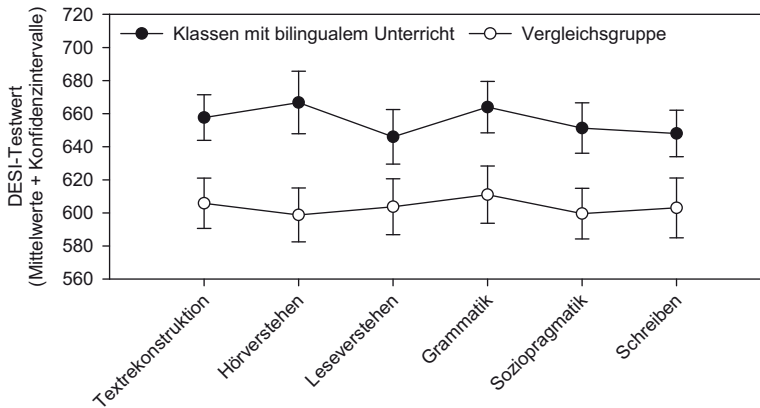


Abbildung 35.1: Mittlere Englisch-Kompetenzen zum Ende der neunten Jahrgangsstufe in Klassen mit bilinguaem Unterricht und in einer hinsichtlich Deutsch Gesamtleistung, Bildungsgang, sozioökonomischem Status, kognitiver Grundfähigkeit, Erstsprache und Geschlecht parallelen Vergleichsgruppe.

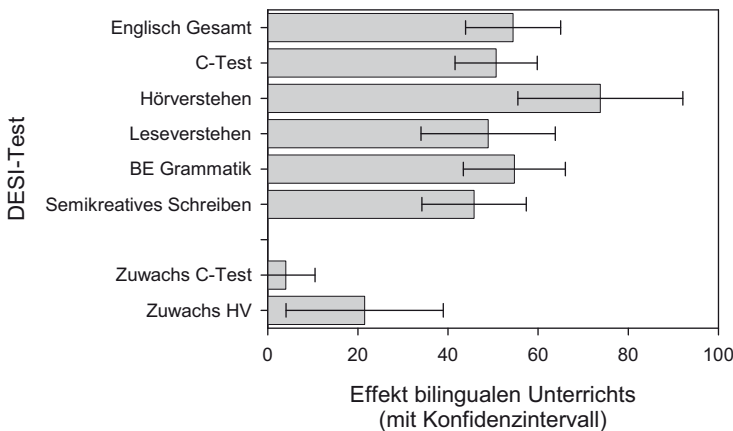


Abbildung 35.2: Effekte bilingualen Unterrichts auf Englisch-Kompetenzen zum Ende der neunten Jahrgangsstufe und auf den Kompetenzzuwachs unter Kontrolle von Deutsch Gesamtleistung, Bildungsgang, sozioökonomischem Status, kognitiver Grundfähigkeit, Erstsprache und Geschlecht.

35.3 Diskussion der Ergebnisse

Die Ergebnisse der Untersuchungen zum bilingualen Sachfachunterricht Englisch in DESI belegen erstmalig in einer großen Stichprobe, dass das Konzept von bilingualem Sachfachunterricht die mit ihm verbundenen Hoffnungen umfassend erfüllt. So werden die Erwartungen in den meisten Bereichen der sprachlichen Kommunikation sowie der Sprachbewusstheit sehr überzeugend erfüllt, wenn nicht sogar übertroffen, da die Schülerinnen und Schüler im Durchschnitt schon in der neunten Klasse ein sprachliches Niveau erreichen, das bei vergleichbaren Klassen ohne ein bilinguales Programm in der Regel erst in der Sekundarstufe II erreicht wird. Die im bilingualen Sachfachunterricht erreichbare Sprachkompetenz sowie die beobachtbare Vertiefung interkultureller (vgl. Kapitel 17) und methodischer Kompetenzen ermöglichen eine Umgestaltung oder Neuausrichtung des Fremdsprachenunterrichts der Sekundarstufe I. Die Schülerinnen und Schüler erleben und verwenden die Fremdsprache Englisch sehr viel stärker als Kommunikationsmedium. Es lassen sich daher schon zu einem früheren Zeitpunkt oder in komplexerer Form kulturell-landeskundliche und interkulturelle Zielsetzungen auf hohem sprachlichen Niveau verwirklichen, die gerade auch in den Lehrplänen der verschiedenen Länder in den Vordergrund gerückt werden.

Literatur

- Bach, G./Niemeier, S. (Hrsg.) (2000): Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven. Frankfurt/M: Lang.
- Bonnet, A. (2004): Chemie im bilingualen Unterricht: Kompetenzerwerb durch Interaktion. Opladen: Leske & Budrich.
- Christ, I. (2002): Bilinguale Module – eine weitere Form des bilingualen Unterrichts? In: Finkbeiner, C. (Hrsg.): Bilingualität und Mehrsprachigkeit: Modelle, Projekte, Ergebnisse. Hannover: Schroedel, S. 14-21.
- Finkbeiner, C. (1995): Englischunterricht in europäischer Dimension. Brockmeyer: Bochum.
- Finkbeiner, C. (Hrsg.) (2002): Bilingualität und Mehrsprachigkeit: Modelle, Projekte, Ergebnisse. Hannover: Schroedel.
- Hallet, W. (1998): The bilingual triangle. Überlegung zu einer Didaktik des bilingualen Sachfachunterrichts. In: Praxis des neusprachlichen Unterrichts 2, S. 115-125.
- Hallet, W. (2005): Bilingualer Unterricht: Fremdsprachig denken, lernen und handeln. In: Der Fremdsprachliche Unterricht Englisch 78, S. 2-8.
- Helbig, B. (2001): Das bilinguale Sachfach Geschichte. Eine empirische Studie zur Arbeit mit französischen (Quellen-)Texten. Tübingen: Stauffenburg.
- Helbig, B./Kleppin, K./Königs, F. G. (Hrsg.) (2002): **Sprachlehrforschung im Wandel. Beiträge zur Erforschung des Lehrens und Lernens von Fremdsprachen.** Festschrift für Karl-Richard Bausch zum 60. Geburtstag. Tübingen: Stauffenburg.
- Krechel, H.-L. (2003): Bilingual Modules – Flexible Formen bilingualen Lehrens und Lernens. In: Wildhage, M./Otten, E. (Hrsg.): Praxis des bilingualen Unterrichts. Berlin: Cornelsen Scriptor, S. 194-215.
- Marsh, D. (Hrsg.) (2002): CLIL/EMILE. The European Dimension. Public Services Contract DG EAC 36 01 Lot 3. September 2002.

- Thürmann, E. (2000): Zwischenbilanz zur Entwicklung der bilingualen Bildungsangebote in Deutschland. Verfügbar unter (09/05): URL: <http://www.learn-line.nrw.de/angebote/berichtsarchiv/pdf/ar0001.pdf>
- Thürmann, E. (2002): Zwischenbilanz zur Entwicklung bilingualer Angebote in Deutschland. In: Helbig, B./Kleppin, K./Königs, F. G. (Hrsg.): Sprachlehrforschung im Wandel. Beiträge zur Erforschung des Lehrens und Lernens von Fremdsprachen. Festschrift für Karl-Richard Bausch zum 60. Geburtstag. Tübingen: Stauffenburg, S. 75-93.
- Vollmer, H. (2000a): Bilingualer Sachfachunterricht als Inhalts- und Sprachenlernen. In: Bach, G./Niemeier, S. (Hrsg.): Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven. Frankfurt/M: Lang, S. 51-73.
- Vollmer, H. (2000b): Förderung des Spracherwerbs im bilingualen Sachfachunterricht. In: Bach, G./Niemeier, S. (Hrsg.): Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven. Frankfurt/M: Lang, S. 139-158.
- Wildhage, M./Otten, E. (Hrsg.) (2003): Praxis des bilingualen Unterrichts. Berlin: Cornelsen Scriptor.
- Zydatiś, W. (2002): Konzeptuelle Grundlagen einer eigenständigen Didaktik des bilingualen Sachfachunterrichts: Forschungsstand und Forschungsprogramm. In: Bonnet, A./Breidbach, S. (Hrsg.): Didaktiken im Dialog. Konzepte des Lehrens und Wege des Lernens im bilingualen Sachfachunterricht. Frankfurt/M.: Lang, S. 31-61.

Die Autorinnen und Autoren

Dr. habil. Bärbel Beck, Didacta Verband e.V., Verband der Bildungswirtschaft,
Rheinstraße 94, 64295 Darmstadt

Svenja Bundt, IEA Data Processing and Research Center, Mexikoring 37, 22297
Hamburg

Kyriaki Chatzivassiliadou, Universität Dortmund, Fakultät Kulturwissenschaften,
Institut für Anglistik und Amerikanistik, Emil-Figge-Straße 50, 44227
Dortmund

Dr. Holger Ehlers, Universität Hamburg, Fachbereich Erziehungswissenschaft,
Institut für Didaktik der Sprachen, Von Melle Park 8, 20146 Hamburg

Prof. Dr. Wolfgang Eichler, Carl-von-Ossietzky-Universität Oldenburg, Fachbereich
Literatur- und Sprachwissenschaften, Institut für Germanistik, Ammerländer
Heerstraße 114 – 118, 26111 Oldenburg

Steffen Gailberger, Johann Wolfgang Goethe-Universität, Institut für Deutsche
Sprache und Literatur I; Grüneburgplatz 1, 60629 Frankfurt am Main

Dr. Kerstin Göbel, Bergische Universität Wuppertal, Fachbereich Bildungs- und
Sozialwissenschaften, Gaustraße 20, 42097 Wuppertal

Jens Gomolka, IEA Data Processing and Research Center, Mexikoring 37, 22297
Hamburg

Dr. Claudia Harsch, Institut zur Qualitätsentwicklung im Bildungswesen an der
Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin

Dr. Johannes Hartig, Deutsches Institut für Internationale Pädagogische Forschung,
Schloßstraße 29, 60486 Frankfurt am Main

Prof. Dr. Andreas Helmke, Universität Koblenz-Landau, Fachbereich Psychologie,
Fortstraße 7, 76829 Landau

Dr. Tuyet Helmke, Universität Koblenz-Landau, Fachbereich Psychologie, Fortstraße
7, 76829 Landau

Dr. Hermann-Günter Hesse, Deutsches Institut für Internationale Pädagogische
Forschung, Schloßstraße 29, 60486 Frankfurt am Main

Silke Hinz, Universität Dortmund, Fakultät Kulturwissenschaften, Institut für Anglis-
tik und Amerikanistik, Emil-Figge-Straße 50, 44227 Dortmund

Nina Jude, Deutsches Institut für Internationale Pädagogische Forschung, Schloß-
straße 29, 60486 Frankfurt am Main

Prof. Dr. Eckhard Klieme, Deutsches Institut für Internationale Pädagogische For-
schung, Schloßstraße 29, 60486 Frankfurt am Main

Michael Krelle, Universität Hildesheim, Fachbereich Kulturwissenschaften und
Ästhetische Kommunikation, Marienburger Platz 22, 31141 Hildesheim

- Prof. Dr. Dr. Dr. h.c. Rainer H. Lehmann, Humboldt-Universität Berlin, Philosophische Fakultät IV, Institut für Allgemeine Pädagogik, Abteilung Empirische Bildungsforschung, Geschwister-Scholl-Straße 7, 10099 Berlin
- Michael Leucht, Institut zur Qualitätsentwicklung im Bildungswesen an der Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin
- Dr. Astrid Neumann, Institut für Schulqualität der Länder Berlin und Brandenburg e.V., Otto-von-Simson-Str. 15, 14195 Berlin
- Prof. Dr. Günter Nold, Universität Dortmund, Fakultät Kulturwissenschaften, Institut für Anglistik und Amerikanistik, Emil-Figge-Straße 50, 44227 Dortmund
- Dominique Rauch, Deutsches Institut für Internationale Pädagogische Forschung, Schloßstraße 29, 60486 Frankfurt am Main
- Prof. Dr. (em.) Hans-Günter Rolff, Institut für Schulentwicklungsforschung, Universität Dortmund, Vogelpothsweg 78, 44227 Dortmund
- Dr. Ernst Rösner, Institut für Schulentwicklungsforschung, Universität Dortmund, Vogelpothsweg 78, 44227 Dortmund
- Henning Rossa, Universität Dortmund, Fakultät Kulturwissenschaften, Institut für Anglistik und Amerikanistik, Emil-Figge-Straße 50, 44227 Dortmund
- Dr. Friedrich-Wilhelm Schrader, Universität Koblenz-Landau, Fachbereich Psychologie, Fortstraße 7, 76829 Landau
- Prof. Dr. Konrad Schröder, Universität Augsburg, Philologisch – Historische Fakultät, Institut für Europäische Kulturgeschichte, Lehrstuhl für Didaktik des Englischen, Universitätsstraße 10, 86135 Augsburg
- Dr. Brigitte Steinert, Deutsches Institut für Internationale Pädagogische Forschung, Schloßstraße 29, 60486 Frankfurt am Main
- Prof. Günther Thomé, Johann Wolfgang Goethe-Universität Frankfurt, FB 10 Neuere Philologien, Grüneburgplatz 1, 60629 Frankfurt am Main
- Jan von der Gathen, Universität Dortmund, Fachbereich Erziehungswissenschaft und Soziologie, Emil-Figge-Str. 91, 44227 Dortmund
- Wolfgang Wagner, Universität Koblenz-Landau, Fachbereich Psychologie, Fortstraße 7, 76829 Landau
- Prof. Dr. Heiner Willenberg, Universität Hamburg, Fachbereich Erziehungswissenschaft, Institut für Didaktik der Sprachen, Von Melle Park 8, 20146 Hamburg



Bärbel Beck / Eckhard Klieme (Hrsg.)
Sprachliche Kompetenzen
Konzepte und Messung.
DESI-Studie (Deutsch Englisch
Schülerleistungen International).
2007. VI, 319 Seiten. Broschiert.
ISBN 978-3-407-25398-7

Was leisten Schülerinnen und Schüler? Welche Kompetenzen kann ihnen der Englischunterricht vermitteln? Wie gelangt man zu Kompetenzmodellen, die Teilbereiche und Niveaus klar abbilden? Mit

welchen Testaufgaben und welchen psychometrischen Methoden kann man fremdsprachliche Kompetenzen messen?

Diese Fragen werden auch in Deutschland verstärkt gestellt, seitdem Leistungsmessung an Schulen und schulübergreifende Vergleichstests zur Regel geworden sind. Bildungsstandards legen fest, welche Kompetenzen Schülerinnen und Schüler erwerben sollen. Tests dienen dazu, die erwarteten Schwierigkeitsabstufungen zu überprüfen und das tatsächlich erreichte Niveau festzustellen.

Mit der Ausschreibung des Projektes »Deutsch Englisch Schülerleistungen International (DESI)« hat die Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland die seit Mitte der 1990er-Jahre laufenden repräsentativen Schulleistungsstudien um neue, zentrale Leistungsbereiche erweitert. Die DESI-Studie erhebt erstmals aussagefähige Daten zu den Fähigkeiten deutscher Schülerinnen und Schüler im rezeptiven und produktiven Gebrauch des Deutschen und Englischen.

BELTZ

Beltz Verlag · Weinheim und Basel · Weitere Infos und Ladenpreis: www.beltz.de